GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 19:59:00 ; Search time 241 Seconds
(without alignments)
2728.433 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 1 | PMS1_HUMAN | P54277 homo sapien |
| 2 | 4742 | 98.5 | 920 | 2 | Q4VAL4_HUMAN | Q4val4 homo sapien |
| 3 | 4693.5 | 97.5 | 931 | 2 | Q5R904_PONPY | Q5r904 pongo pygma |
| 4 | 4584.5 | 95.3 | 893 | 2 | Q5FBZ3_HUMAN | Q5fbz3 homo sapien |
| 5 | 3889 | 80.8 | 770 | 2 | Q5FBZ8_HUMAN | Q5fbz8 homo sapien |
| 6 | 3579.5 | 74.4 | 917 | 2 | Q8K119_MOUSE | Q8k119 mus musculu |
| 7 | 3543.5 | 73.6 | 919 | 2 | Q6P7D0_RAT | Q6p7d0 rattus norv |
| 8 | 3351 | 69.6 | 669 | 2 | Q4VAL5_HUMAN | Q4val5 homo sapien |
| 9 | 3349 | 69.6 | 667 | 2 | Q5FBZ9_HUMAN | Q5fbz9 homo sapien |
| 10 | 2779.5 | 57.8 | 555 | 2 | Q5FBZ6_HUMAN | Q5fbz6 homo sapien |
| 11 | 2720 | 56.5 | 916 | 2 | Q5ZKT5_CHICK | Q5zkt5 gallus gall |
| 12 | 2319 | 48.2 | 925 | 2 | Q7ZXV9_XENLA | Q7zxv9 xenopus lae |
| 13 | 2213.5 | 46.0 | 928 | 2 | Q5FVX9_XENTR | Q5fvx9 xenopus tro |
| 14 | 2046.5 | 42.5 | 896 | 2 | Q8JFR9_BRARE | Q8jfr9 brachydanio |
| 15 | 1835.5 | 38.1 | 854 | 2 | Q4RTJ3_TETNG | Q4rtj3 tetraodon n |
| 16 | 1177 | 24.5 | 234 | 2 | Q5FBZ2_HUMAN | Q5fbz2 homo sapien |
| 17 | 1132.5 | 23.5 | 372 | 2 | Q7SXD5_BRARE | Q7sxd5 brachydanio |
| 18 | 977 | 20.3 | 248 | 2 | Q5FBZ4_HUMAN | Q5fbz4 homo sapien |
| 19 | 976 | 20.3 | 196 | 2 | Q5FBZ5_HUMAN | Q5fbz5 homo sapien |
| 20 | 975 | 20.3 | 195 | 2 | Q5FBZ1_HUMAN | Q5fbz1 homo sapien |
| 21 | 830 | 17.2 | 163 | 2 | Q68DF0_HUMAN | Q68df0 homo sapien |
| 22 | 719 | 14.9 | 194 | 2 | Q8JFW5_BRARE | Q8jfw5 brachydanio |
| 23 | 708 | 14.7 | 166 | 2 | Q96HL0_HUMAN | Q96hl0 homo sapien |
| 24 | 702 | 14.6 | 165 | 2 | Q5XG96_HUMAN | Q5xg96 homo sapien |
| 25 | 659.5 | 13.7 | 880 | 2 | Q7QIY1_ANOGA | Q7qiy1 anopheles g |
| 26 | 654 | 13.6 | 1022 | 2 | Q54QA0_DICDI | Q54qa0 dictyosteli |
| 27 | 639 | 13.3 | 143 | 2 | Q8BLI9_MOUSE | Q8bli9 mus musculu |
| 28 | 616.5 | 12.8 | 895 | 2 | Q8T9C0_DROME | Q8t9c0 drosophila |
| 29 | 615.5 | 12.8 | 899 | 2 | Q9V7B6_DROME | Q9v7b6 drosophila |
| 30 | 607 | 12.6 | 923 | 2 | Q941I6_ARATH | Q941i6 arabidopsis |
| 31 | 595.5 | 12.4 | 893 | 2 | O76417_DROME | O76417 drosophila |
| 32 | 587.5 | 12.2 | 805 | 2 | Q9TVL8_CAEEL | Q9tvl8 caenorhabdi |
| 33 | 584 | 12.1 | 871 | 2 | Q5ZJ94_CHICK | Q5zj94 gallus gall |
| 34 | 581.5 | 12.1 | 903 | 2 | Q755U7_ASHGO | Q755u7 ashbya goss |
| 35 | 572.5 | 11.9 | 797 | 2 | Q60M36_CAEBR | Q60m36 caenorhabdi |
| 36 | 569 | 11.8 | 862 | 1 | PMS2_HUMAN | P54278 homo sapien |
| 37 | 569 | 11.8 | 862 | 2 | Q75MR2_HUMAN | Q75mr2 homo sapien |
| 38 | 569 | 11.8 | 879 | 2 | Q8N5Q6_HUMAN | Q8n5q6 homo sapien |
| 39 | 568 | 11.8 | 862 | 2 | Q52LH6_HUMAN | Q52lh6 homo sapien |
| 40 | 564.5 | 11.7 | 907 | 2 | Q6FPA0_CANGA | Q6fpa0 candida gla |
| 41 | 563 | 11.7 | 923 | 2 | Q69L72_ORYSA | Q69l72 oryza sativ |
| 42 | 561.5 | 11.7 | 893 | 2 | Q6C6B8_YARLI | Q6c6b8 yarrowia li |
| 43 | 561 | 11.7 | 908 | 2 | Q8TG50_YEAST | Q8tg50 saccharomyc |
| 44 | 559.5 | 11.6 | 1094 | 2 | Q4XWC3_PLACH | Q4xwc3 plasmodium |
| 45 | 558 | 11.6 | 908 | 2 | Q8TG48_YEAST | Q8tg48 saccharomyc |

ALIGNMENTS

RESULT 1
PMS1_HUMAN
ID PMS1_HUMAN STANDARD; PRT; 932 AA.
AC P54277;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE PMS1 protein homolog 1 (DNA mismatch repair protein PMS1).
GN Name=PMS1; Synonyms=PMSL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Gall bladder;
RX MEDLINE=94352394; PubMed=8072530; DOI=10.1038/371075a0;
RA Nicolaides N.C., Papadopoulos N., Liu B., Wei Y.-F., Carter K.C.,
RA Ruben S.M., Rosen C.A., Haseltine W.H., Fleischmann R.D., Fraser C.M.,
RA Adams M.D., Venter J.C., Dunlop M.G., Hamilton S.R., Petersen G.M.,
RA de la Chapelle A., Vogelstein B., Kinzler K.W.;
RT "Mutations of two PMS homologues in hereditary nonpolyposis colon
RT cancer.";
RL Nature 371:75-80(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS GLN-27; LYS-202;
RP ARG-501; SER-632; ASP-720 AND HIS-793.
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT "NIEHS-SNPs, environmental genome project, NIEHS ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP VARIANTS HNPCC3 THR-394 AND ARG-501.
RX MEDLINE=99408236; PubMed=10480359; DOI=10.1007/s004390051067;
RA Wang Q., Lasset C., Desseigne F., Saurin J.-C., Maugard C.,
RA Navarro C., Ruano E., Descos L., Trillet-Lenoir V., Bosset J.-F.,
RA Puisieux A.;
RT "Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and
RT hMSH6 genes in 75 French kindreds with nonpolyposis colorectal
RT cancer.";
RL Hum. Genet. 105:79-85(1999).
CC -!- FUNCTION: Probably involved in the repair of mismatches in DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DISEASE: Defects in PMS1 are the cause of hereditary non-polyposis
CC colorectal cancer type 3 (HNPCC3) [MIM:600258]. Mutations in more
CC than one gene locus can be involved alone or in combination in the
CC production of the HNPCC phenotype (also called Lynch syndrome).
CC Most families with clinically recognized HNPCC have mutations in
CC either MLH1 or MSH2 genes. HNPCC is an autosomal, dominantly
CC inherited disease associated with marked increase in cancer
CC susceptibility. It is characterized by a familial predisposition

```
CC       to early onset colorectal carcinoma (CRC) and extra-colonic
CC       cancers of the gastrointestinal, urological and female
CC       reproductive tracts. HNPCC is reported to be the most common form
CC       of inherited colorectal cancer in the Western world, and accounts
CC       for 15% of all colon cancers. Cancers in HNPCC originate within
CC       benign neoplastic polyps termed adenomas. Clinically, HNPCC is
CC       often divided into two subgroups. Type I: hereditary
CC       predisposition to colorectal cancer, a young age of onset, and
CC       carcinoma observed in the proximal colon. Type II: patients have
CC       an increased risk for cancers in certain tissues such as the
CC       uterus, ovary, breast, stomach, small intestine, skin, and larynx
CC       in addition to the colon. Diagnosis of classical HNPCC is based on
CC       the Amsterdam criteria: 3 or more relatives affected by colorectal
CC       cancer, one a first degree relative of the other two; 2 or more
CC       generation affected; 1 or more colorectal cancers presenting
CC       before 50 years of age; exclusion of hereditary polyposis
CC       syndromes. The term "suspected HNPCC" or "incomplete HNPCC" can be
CC       used to describe families who do not or only partially fulfill the
CC       Amsterdam criteria, but in whom a genetic basis for colon cancer
CC       is strongly suspected.
CC   -!- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB family.
CC   -!- SIMILARITY: Contains 1 HMG box DNA-binding domain.
CC   -!- DATABASE: NAME=Hereditary non-polyposis colorectal cancer db;
CC       WWW="http://www.nfdht.nl/".
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U13695; AAA63922.1; -; Genomic_DNA.
DR   EMBL; AY267352; AAO89079.1; -; Genomic_DNA.
DR   PIR; S47597; S47597.
DR   HSSP; P54278; 1H7S.
DR   Ensembl; ENSG00000064933; Homo sapiens.
DR   HGNC; HGNC:9121; PMS1.
DR   MIM; 600258; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003677; F:DNA binding; TAS.
DR   GO; GO:0006298; P:mismatch repair; TAS.
DR   InterPro; IPR003594; ATP_bd_ATPase.
DR   InterPro; IPR002099; DNA_mis_repair.
DR   InterPro; IPR000910; HMG_12_box.
DR   PANTHER; PTHR10073; DNA_mis_repair; 1.
DR   Pfam; PF01119; DNA_mis_repair; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00505; HMG_box; 1.
DR   TIGRFAMs; TIGR00585; mutl; 1.
DR   PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 1.
KW   Anti-oncogene; Cell cycle; Disease mutation; DNA damage; DNA repair;
KW   Hereditary nonpolyposis colorectal cancer; Nuclear protein;
KW   Polymorphism.
FT   DNA_BIND     571    639       HMG box.
FT   VARIANT       27     27       E -> Q (in dbSNP:5742973).
FT                                 /FTId=VAR_019166.
FT   VARIANT      202    202       R -> K (in dbSNP:2066459).
FT                                 /FTId=VAR_014877.
FT   VARIANT      394    394       M -> T (in incomplete HNPCC3;
FT                                 dbSNP:1145231).
FT                                 /FTId=VAR_012967.
FT   VARIANT      501    501       G -> R (in incomplete HNPCC3;
FT                                 dbSNP:1145232).
FT                                 /FTId=VAR_012968.
FT   VARIANT      632    632       N -> S (in dbSNP:2066456).
FT                                 /FTId=VAR_014878.
FT   VARIANT      720    720       E -> D (in dbSNP:2066455).
FT                                 /FTId=VAR_014879.
FT   VARIANT      793    793       Y -> H (in dbSNP:1145234).
FT                                 /FTId=VAR_014880.
SQ   SEQUENCE   932 AA;  105830 MW;  EC4F402937B616DF CRC64;
```

```
  Query Match           100.0%;  Score 4812;  DB 1;  Length 932;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-214;
  Matches  932;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932


RESULT 2
Q4VAL4_HUMAN
ID   Q4VAL4_HUMAN PRELIMINARY;      PRT;   920 AA.
AC   Q4VAL4;
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 19:59:00 ; Search time 241 Seconds
(without alignments)
2728.433 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 1 | PMS1_HUMAN | P54277 homo sapien |
| 2 | 4742 | 98.5 | 920 | 2 | Q4VAL4_HUMAN | Q4val4 homo sapien |
| 3 | 4693.5 | 97.5 | 931 | 2 | Q5R904_PONPY | Q5r904 pongo pygma |
| 4 | 4584.5 | 95.3 | 893 | 2 | Q5FBZ3_HUMAN | Q5fbz3 homo sapien |
| 5 | 3889 | 80.8 | 770 | 2 | Q5FBZ8_HUMAN | Q5fbz8 homo sapien |
| 6 | 3579.5 | 74.4 | 917 | 2 | Q8K119_MOUSE | Q8k119 mus musculu |
| 7 | 3543.5 | 73.6 | 919 | 2 | Q6P7D0_RAT | Q6p7d0 rattus norv |
| 8 | 3351 | 69.6 | 669 | 2 | Q4VAL5_HUMAN | Q4val5 homo sapien |
| 9 | 3349 | 69.6 | 667 | 2 | Q5FBZ9_HUMAN | Q5fbz9 homo sapien |
| 10 | 2779.5 | 57.8 | 555 | 2 | Q5FBZ6_HUMAN | Q5fbz6 homo sapien |
| 11 | 2720 | 56.5 | 916 | 2 | Q5ZKT5_CHICK | Q5zkt5 gallus gall |
| 12 | 2319 | 48.2 | 925 | 2 | Q7ZXV9_XENLA | Q7zxv9 xenopus lae |
| 13 | 2213.5 | 46.0 | 928 | 2 | Q5FVX9_XENTR | Q5fvx9 xenopus tro |
| 14 | 2046.5 | 42.5 | 896 | 2 | Q8JFR9_BRARE | Q8jfr9 brachydanio |
| 15 | 1835.5 | 38.1 | 854 | 2 | Q4RTJ3_TETNG | Q4rtj3 tetraodon n |
| 16 | 1177 | 24.5 | 234 | 2 | Q5FBZ2_HUMAN | Q5fbz2 homo sapien |
| 17 | 1132.5 | 23.5 | 372 | 2 | Q7SXD5_BRARE | Q7sxd5 brachydanio |
| 18 | 977 | 20.3 | 248 | 2 | Q5FBZ4_HUMAN | Q5fbz4 homo sapien |
| 19 | 976 | 20.3 | 196 | 2 | Q5FBZ5_HUMAN | Q5fbz5 homo sapien |
| 20 | 975 | 20.3 | 195 | 2 | Q5FBZ1_HUMAN | Q5fbz1 homo sapien |
| 21 | 830 | 17.2 | 163 | 2 | Q68DF0_HUMAN | Q68df0 homo sapien |
| 22 | 719 | 14.9 | 194 | 2 | Q8JFW5_BRARE | Q8jfw5 brachydanio |
| 23 | 708 | 14.7 | 166 | 2 | Q96HL0_HUMAN | Q96hl0 homo sapien |
| 24 | 702 | 14.6 | 165 | 2 | Q5XG96_HUMAN | Q5xg96 homo sapien |
| 25 | 659.5 | 13.7 | 880 | 2 | Q7QIY1_ANOGA | Q7qiy1 anopheles g |
| 26 | 654 | 13.6 | 1022 | 2 | Q54QA0_DICDI | Q54qa0 dictyosteli |
| 27 | 639 | 13.3 | 143 | 2 | Q8BLI9_MOUSE | Q8bli9 mus musculu |
| 28 | 616.5 | 12.8 | 895 | 2 | Q8T9C0_DROME | Q8t9c0 drosophila |
| 29 | 615.5 | 12.8 | 899 | 2 | Q9V7B6_DROME | Q9v7b6 drosophila |
| 30 | 607 | 12.6 | 923 | 2 | Q941I6_ARATH | Q941i6 arabidopsis |
| 31 | 595.5 | 12.4 | 893 | 2 | O76417_DROME | O76417 drosophila |
| 32 | 587.5 | 12.2 | 805 | 2 | Q9TVL8_CAEEL | Q9tvl8 caenorhabdi |
| 33 | 584 | 12.1 | 871 | 2 | Q5ZJ94_CHICK | Q5zj94 gallus gall |
| 34 | 581.5 | 12.1 | 903 | 2 | Q755U7_ASHGO | Q755u7 ashbya goss |
| 35 | 572.5 | 11.9 | 797 | 2 | Q60M36_CAEBR | Q60m36 caenorhabdi |
| 36 | 569 | 11.8 | 862 | 1 | PMS2_HUMAN | P54278 homo sapien |
| 37 | 569 | 11.8 | 862 | 2 | Q75MR2_HUMAN | Q75mr2 homo sapien |
| 38 | 569 | 11.8 | 879 | 2 | Q8N5Q6_HUMAN | Q8n5q6 homo sapien |
| 39 | 568 | 11.8 | 862 | 2 | Q52LH6_HUMAN | Q52lh6 homo sapien |
| 40 | 564.5 | 11.7 | 907 | 2 | Q6FPA0_CANGA | Q6fpa0 candida gla |
| 41 | 563 | 11.7 | 923 | 2 | Q69L72_ORYSA | Q69l72 oryza sativ |
| 42 | 561.5 | 11.7 | 893 | 2 | Q6C6B8_YARLI | Q6c6b8 yarrowia li |
| 43 | 561 | 11.7 | 908 | 2 | Q8TG50_YEAST | Q8tg50 saccharomyc |
| 44 | 559.5 | 11.6 | 1094 | 2 | Q4XWC3_PLACH | Q4xwc3 plasmodium |
| 45 | 558 | 11.6 | 908 | 2 | Q8TG48_YEAST | Q8tg48 saccharomyc |

ALIGNMENTS

```
RESULT 1
PMS1_HUMAN
ID   PMS1_HUMAN      STANDARD;      PRT;   932 AA.
AC   P54277;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   PMS1 protein homolog 1 (DNA mismatch repair protein PMS1).
GN   Name=PMS1; Synonyms=PMSL1;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Gall bladder;
RX   MEDLINE=94352394; PubMed=8072530; DOI=10.1038/371075a0;
RA   Nicolaides N.C., Papadopoulos N., Liu B., Wei Y.-F., Carter K.C.,
RA   Ruben S.M., Rosen C.A., Haseltine W.H., Fleischmann R.D., Fraser C.M.,
RA   Adams M.D., Venter J.C., Dunlop M.G., Hamilton S.R., Petersen G.M.,
RA   de la Chapelle A., Vogelstein B., Kinzler K.W.;
RT   "Mutations of two PMS homologues in hereditary nonpolyposis colon
RT   cancer.";
RL   Nature 371:75-80(1994).
RN   [2]
RP   NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS GLN-27; LYS-202;
RP   ARG-501; SER-632; ASP-720 AND HIS-793.
RA   Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
RA   Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA   Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT   "NIEHS-SNPs, environmental genome project, NIEHS ES15478, Department
RT   of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   VARIANTS HNPCC3 THR-394 AND ARG-501.
RX   MEDLINE=99408236; PubMed=10480359; DOI=10.1007/s004390051067;
RA   Wang Q., Lasset C., Desseigne F., Saurin J.-C., Maugard C.,
RA   Navarro C., Ruano E., Descos L., Trillet-Lenoir V., Bosset J.-F.,
RA   Puisieux A.;
RT   "Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and
RT   hMSH6 genes in 75 French kindreds with nonpolyposis colorectal
RT   cancer.";
RL   Hum. Genet. 105:79-85(1999).
CC   -!- FUNCTION: Probably involved in the repair of mismatches in DNA.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC   -!- DISEASE: Defects in PMS1 are the cause of hereditary non-polyposis
CC       colorectal cancer type 3 (HNPCC3) [MIM:600258]. Mutations in more
CC       than one gene locus can be involved alone or in combination in the
CC       production of the HNPCC phenotype (also called Lynch syndrome).
CC       Most families with clinically recognized HNPCC have mutations in
CC       either MLH1 or MSH2 genes. HNPCC is an autosomal, dominantly
CC       inherited disease associated with marked increase in cancer
CC       susceptibility. It is characterized by a familial predisposition
```

```
CC       to early onset colorectal carcinoma (CRC) and extra-colonic
CC       cancers of the gastrointestinal, urological and female
CC       reproductive tracts. HNPCC is reported to be the most common form
CC       of inherited colorectal cancer in the Western world, and accounts
CC       for 15% of all colon cancers. Cancers in HNPCC originate within
CC       benign neoplastic polyps termed adenomas. Clinically, HNPCC is
CC       often divided into two subgroups. Type I: hereditary
CC       predisposition to colorectal cancer, a young age of onset, and
CC       carcinoma observed in the proximal colon. Type II: patients have
CC       an increased risk for cancers in certain tissues such as the
CC       uterus, ovary, breast, stomach, small intestine, skin, and larynx
CC       in addition to the colon. Diagnosis of classical HNPCC is based on
CC       the Amsterdam criteria: 3 or more relatives affected by colorectal
CC       cancer, one a first degree relative of the other two; 2 or more
CC       generation affected; 1 or more colorectal cancers presenting
CC       before 50 years of age; exclusion of hereditary polyposis
CC       syndromes. The term "suspected HNPCC" or "incomplete HNPCC" can be
CC       used to describe families who do not or only partially fulfill the
CC       Amsterdam criteria, but in whom a genetic basis for colon cancer
CC       is strongly suspected.
CC   -!- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB family.
CC   -!- SIMILARITY: Contains 1 HMG box DNA-binding domain.
CC   -!- DATABASE: NAME=Hereditary non-polyposis colorectal cancer db;
CC       WWW="http://www.nfdht.nl/".
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U13695; AAA63922.1; -; Genomic_DNA.
DR   EMBL; AY267352; AAO89079.1; -; Genomic_DNA.
DR   PIR; S47597; S47597.
DR   HSSP; P54278; 1H7S.
DR   Ensembl; ENSG00000064933; Homo sapiens.
DR   HGNC; HGNC:9121; PMS1.
DR   MIM; 600258; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003677; F:DNA binding; TAS.
DR   GO; GO:0006298; P:mismatch repair; TAS.
DR   InterPro; IPR003594; ATP_bd_ATPase.
DR   InterPro; IPR002099; DNA_mis_repair.
DR   InterPro; IPR000910; HMG_12_box.
DR   PANTHER; PTHR10073; DNA_mis_repair; 1.
DR   Pfam; PF01119; DNA_mis_repair; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00505; HMG_box; 1.
DR   TIGRFAMs; TIGR00585; mutl; 1.
DR   PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 1.
KW   Anti-oncogene; Cell cycle; Disease mutation; DNA damage; DNA repair;
KW   Hereditary nonpolyposis colorectal cancer; Nuclear protein;
KW   Polymorphism.
FT   DNA_BIND     571     639       HMG box.
FT   VARIANT       27      27       E -> Q (in dbSNP:5742973).
FT                                  /FTId=VAR_019166.
FT   VARIANT      202     202       R -> K (in dbSNP:2066459).
FT                                  /FTId=VAR_014877.
FT   VARIANT      394     394       M -> T (in incomplete HNPCC3;
FT                                  dbSNP:1145231).
FT                                  /FTId=VAR_012967.
FT   VARIANT      501     501       G -> R (in incomplete HNPCC3;
FT                                  dbSNP:1145232).
FT                                  /FTId=VAR_012968.
FT   VARIANT      632     632       N -> S (in dbSNP:2066456).
FT                                  /FTId=VAR_014878.
FT   VARIANT      720     720       E -> D (in dbSNP:2066455).
FT                                  /FTId=VAR_014879.
FT   VARIANT      793     793       Y -> H (in dbSNP:1145234).
FT                                  /FTId=VAR_014880.
SQ   SEQUENCE   932 AA;  105830 MW;  EC4F402937B616DF CRC64;
```

```
  Query Match           100.0%;  Score 4812;  DB 1;  Length 932;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-214;
  Matches  932;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932


RESULT 2
Q4VAL4_HUMAN
ID   Q4VAL4_HUMAN PRELIMINARY;      PRT;   920 AA.
AC   Q4VAL4;
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:24:27 ; Search time 46 Seconds
(without alignments)
1949.435 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 932
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 932 | 100.0 | 932 | 2 | S47597 | mutL protein homol |
| 2 | 16 | 1.7 | 580 | 2 | A72032 | DNA mismatch repai |
| 3 | 16 | 1.7 | 580 | 2 | B86592 | DNA mismatch repai |
| 4 | 15 | 1.6 | 576 | 2 | A71497 | probable DNA misma |
| 5 | 12 | 1.3 | 779 | 2 | T01304 | hypothetical prote |
| 6 | 11 | 1.2 | 684 | 2 | T50317 | probable DNA misma |
| 7 | 9 | 1.0 | 98 | 2 | JC2403 | PMS8 homolog misma |
| 8 | 9 | 1.0 | 159 | 2 | JC2401 | PMS6 homolog misma |
| 9 | 9 | 1.0 | 161 | 2 | JC2402 | PMS7 homolog misma |
| 10 | 9 | 1.0 | 186 | 2 | JC2400 | PMS5 homolog misma |
| 11 | 9 | 1.0 | 249 | 2 | D96691 | hypothetical prote |
| 12 | 9 | 1.0 | 252 | 2 | JC2399 | PMS4 homolog misma |
| 13 | 9 | 1.0 | 256 | 2 | JC2398 | PMS3 homolog misma |
| 14 | 9 | 1.0 | 425 | 2 | D70436 | DNA mismatch repai |
| 15 | 9 | 1.0 | 516 | 2 | H72427 | DNA mismatch repai |
| 16 | 9 | 1.0 | 530 | 2 | C90248 | DNA topoisomerase |
| 17 | 9 | 1.0 | 576 | 2 | G81657 | DNA mismatch repai |
| 18 | 9 | 1.0 | 615 | 2 | PH0853 | methyl-directed mi |
| 19 | 9 | 1.0 | 615 | 2 | B91272 | enzyme in methyl-d |
| 20 | 9 | 1.0 | 615 | 2 | B86113 | enzyme in methyl-d |
| 21 | 9 | 1.0 | 618 | 2 | AG1048 | DNA mismatch repai |
| 22 | 9 | 1.0 | 618 | 2 | A33588 | mismatch repair pr |
| 23 | 9 | 1.0 | 635 | 2 | AC0046 | DNA mismatch repai |
| 24 | 9 | 1.0 | 653 | 2 | A82334 | DNA mismatch repai |
| 25 | 9 | 1.0 | 756 | 2 | S43085 | DNA mismatch repai |
| 26 | 9 | 1.0 | 769 | 2 | S54525 | mismatch repair pr |
| 27 | 9 | 1.0 | 779 | 2 | T25389 | hypothetical prote |
| 28 | 9 | 1.0 | 862 | 2 | S47598 | mutL protein homol |
| 29 | 8 | 0.9 | 114 | 2 | T16365 | hypothetical prote |
| 30 | 8 | 0.9 | 116 | 2 | G84032 | hypothetical prote |
| 31 | 8 | 0.9 | 213 | 2 | D69409 | conserved hypothet |
| 32 | 8 | 0.9 | 259 | 2 | A86409 | hypothetical prote |
| 33 | 8 | 0.9 | 305 | 2 | B84274 | heme biosynthesis |
| 34 | 8 | 0.9 | 336 | 2 | A81409 | probable periplasm |
| 35 | 8 | 0.9 | 349 | 2 | H64042 | signal peptidase I |
| 36 | 8 | 0.9 | 392 | 2 | F96937 | cell wall-associat |
| 37 | 8 | 0.9 | 393 | 2 | AG0091 | probable flagellar |
| 38 | 8 | 0.9 | 446 | 2 | D71418 | hypothetical prote |
| 39 | 8 | 0.9 | 474 | 2 | B86221 | hypothetical prote |
| 40 | 8 | 0.9 | 492 | 2 | T00433 | fumarate hydratase |
| 41 | 8 | 0.9 | 532 | 2 | H83993 | two-component sens |
| 42 | 8 | 0.9 | 563 | 2 | AH2187 | DNA mismatch repai |
| 43 | 8 | 0.9 | 584 | 2 | JQ1229 | cellulase (EC 3.2. |
| 44 | 8 | 0.9 | 629 | 2 | E64046 | mismatch repair pr |
| 45 | 8 | 0.9 | 659 | 2 | E84176 | DNA mismatch repai |

ALIGNMENTS

RESULT 1
S47597
mutL protein homolog - human
C;Species: Homo sapiens (man)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47597
R;Nicolaides, N.C.; Papadopoulos, N.; Liu, B.; Wei, Y.F.; Carter, K.C.; Ruben, S.M.; R
S.R.; Petersen, G.M.; de la Chapelle, A.; Vogelstein, B.; Kinzler, K.W.
Nature 371, 75-80, 1994
A;Title: Mutations of two PMS homologues in hereditary nonpolyposis colon cancer.
A;Reference number: S47597; MUID:94352394; PMID:8072530
A;Accession: S47597
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-932 <NIC>
A;Cross-references: UNIPROT:P54277; UNIPARC:UPI00000405F5; EMBL:U13695; NID:g535512; P
C;Genetics:
A;Gene: GDB:PMS1; PMSL1
A;Cross-references: GDB:386403; OMIM:600258
A;Map position: 2q31-2q33
F;571-643/Domain: HMG box homology <HMG1>

Query Match 100.0%; Score 932; DB 2; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
A72032
DNA mismatch repair protein mutL CP1059 [imported] - Chlamydophila pneumoniae (strains C
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: A72032; F81506
R;Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A;Title: Comparative genomes of Clamydia pneumoniae and C. trachomatis.
A;Reference number: A72000; MUID:99206606; PMID:10192388
A;Accession: A72032
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <ARN>
A;Cross-references: UNIPROT:Q9Z794; UNIPARC:UPI000012FA3C; GB:AE001662; GB:AE001363; NID
A;Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: F81506
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <REA>
A;Cross-references: UNIPARC:UPI000012FA3C; GB:AE002263; GB:AE002161; NID:g7189971; PIDN:
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: mutL; CP1059
C;Superfamily: mismatch repair protein hexB

```
  Query Match             1.7%;  Score 16;  DB 2;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 VSVVKELIENSLDAGA 37
              ||||||||||||||||
Db         28 VSVVKELIENSLDAGA 43
```

RESULT 3
B86592
DNA mismatch repair [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86592
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349; PMID:10871362
A;Accession: B86592
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <STO>
A;Cross-references: UNIPROT:Q9Z794; UNIPARC:UPI000012FA3C; GB:BA000008; NID:g8979186;
A;Experimental source: strain J138
C;Genetics:
A;Gene: mutL
C;Superfamily: mismatch repair protein hexB

```
  Query Match             1.7%;  Score 16;  DB 2;  Length 580;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         22 VSVVKELIENSLDAGA 37
              ||||||||||||||||
Db         28 VSVVKELIENSLDAGA 43
```

RESULT 4
A71497
probable DNA mismatch repair - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C;Accession: A71497
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchel
Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tr
A;Reference number: A71570; MUID:99000809; PMID:9784136
A;Accession: A71497
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-576 <ARN>
A;Cross-references: UNIPROT:O84579; UNIPARC:UPI000012FA3E; GB:AE001328; GB:AE001273; N
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: mutL
C;Superfamily: mismatch repair protein hexB

```
  Query Match             1.6%;  Score 15;  DB 2;  Length 576;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         23 SVVKELIENSLDAGA 37
              |||||||||||||||
Db         30 SVVKELIENSLDAGA 44
```

RESULT 5
T01304
hypothetical protein T14P8.6 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C;Accession: T01304
R;Kalicki, J.; Elliott, G.; Cloud, J.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of A. thaliana T14P8.

```
A;Reference number: Z14290
A;Accession: T01304
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-779 <KAL>
A;Cross-references: UNIPROT:O81287; UNIPARC:UPI000004068B; EMBL:AF069298; NID:g3193282;
A;Experimental source: cultivar Columbia
C;Genetics:
A;Map position: 4
A;Introns: 94/3; 202/3; 254/3; 562/3; 585/3; 610/2; 632/2; 667/1; 727/3
A;Note: T14P8.6
C;Superfamily: DNA mismatch repair protein

  Query Match             1.3%;  Score 12;  DB 2;  Length 779;
  Best Local Similarity   100.0%;  Pred. No. 0.0021;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         86 NLTTYGFRGEAL 97
              ||||||||||||
Db        113 NLTTYGFRGEAL 124


RESULT 6
T50317
probable DNA mismatch repair protein, MLH1 homolog [imported] - fission yeast (Schizosac
C;Species: Schizosaccharomyces pombe
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
C;Accession: T50317
R;McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.; Cadieu, E.; Lelaure, V.; Galibert, F
submitted to the EMBL Data Library, January 2000
A;Reference number: Z25061
A;Accession: T50317
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-684 <MCD>
A;Cross-references: UNIPROT:Q9P7W6; UNIPARC:UPI000006AC8D; EMBL:AL136536; PIDN:CAB66448.
A;Experimental source: strain 972h(-); cosmid c1703
C;Genetics:
A;Gene: SPDB:SPBC1703.04
A;Map position: 2
A;Introns: 24/3; 70/3; 128/2
C;Superfamily: DNA mismatch repair protein, Mlh1 type

  Query Match             1.2%;  Score 11;  DB 2;  Length 684;
  Best Local Similarity   100.0%;  Pred. No. 0.02;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         26 KELIENSLDAG 36
              ||||||||||||
Db         34 KELIENSLDAG 44


RESULT 7
JC2403
PMS8 homolog mismatch repair protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: JC2403
R;Horii, A.; Han, H.J.; Sasaki, S.; Shimada, M.; Nakamura, Y.
Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994
A;Title: Cloning, characterization and chromosomal assignment of the human genes homolog
A;Reference number: JC2398; MUID:95071462; PMID:7980603
A;Accession: JC2403
A;Molecule type: DNA
A;Residues: 1-98 <HOR>
A;Cross-references: UNIPROT:Q16590; UNIPARC:UPI0000073FED; DDBJ:D38440; NID:g600595; PID
C;Genetics:
A;Gene: GDB:PMS2L6; PMS8
A;Cross-references: GDB:437147
A;Map position: 7q11.23-7q22
C;Keywords: DNA repair

  Query Match             1.0%;  Score 9;  DB 2;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 0.44;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 ENSLDAGAT 38
              |||||||||
Db         48 ENSLDAGAT 56


RESULT 8
JC2401
PMS6 homolog mismatch repair protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: JC2401
R;Horii, A.; Han, H.J.; Sasaki, S.; Shimada, M.; Nakamura, Y.
Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994
A;Title: Cloning, characterization and chromosomal assignment of the human genes homol
A;Reference number: JC2398; MUID:95071462; PMID:7980603
A;Accession: JC2401
A;Molecule type: DNA
A;Residues: 1-159 <HOR>
A;Cross-references: UNIPROT:Q16544; UNIPARC:UPI0000073EE1; DDBJ:D38438; NID:g600593; P
C;Genetics:
A;Gene: GDB:PMS2L4; PMS6
A;Cross-references: GDB:437145
A;Map position: 7q11.23-7q22
C;Keywords: DNA repair

  Query Match             1.0%;  Score 9;  DB 2;  Length 159;
  Best Local Similarity   100.0%;  Pred. No. 0.68;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 ENSLDAGAT 38
              |||||||||
Db         49 ENSLDAGAT 57


RESULT 9
JC2402
PMS7 homolog mismatch repair protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: JC2402
R;Horii, A.; Han, H.J.; Sasaki, S.; Shimada, M.; Nakamura, Y.
Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994
A;Title: Cloning, characterization and chromosomal assignment of the human genes homol
A;Reference number: JC2398; MUID:95071462; PMID:7980603
A;Accession: JC2402
A;Molecule type: DNA
A;Residues: 1-161 <HOR>
A;Cross-references: UNIPROT:Q16673; UNIPARC:UPI000017C30B; DDBJ:D38439
C;Genetics:
A;Gene: GDB:PMS2L5; PMS7
A;Cross-references: GDB:437146
A;Map position: 7q11.23-7q22
C;Keywords: DNA repair

  Query Match             1.0%;  Score 9;  DB 2;  Length 161;
  Best Local Similarity   100.0%;  Pred. No. 0.69;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         30 ENSLDAGAT 38
              |||||||||
Db         33 ENSLDAGAT 41


RESULT 10
JC2400
PMS5 homolog mismatch repair protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: JC2400
R;Horii, A.; Han, H.J.; Sasaki, S.; Shimada, M.; Nakamura, Y.
```

Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994
A;Title: Cloning, characterization and chromosomal assignment of the human genes homolog
A;Reference number: JC2398; MUID:95071462; PMID:7980603
A;Accession: JC2400
A;Molecule type: DNA
A;Residues: 1-186 <HOR>
A;Cross-references: UNIPROT:Q16603; UNIPARC:UPI0000073EE3; DDBJ:D38437; NID:g600592; PID
C;Genetics:
A;Gene: GDB:PMS2L3; PMS5
A;Cross-references: GDB:437144
A;Map position: 7q11.23-7q22
C;Keywords: DNA repair

  Query Match 1.0%; Score 9; DB 2; Length 186;
  Best Local Similarity 100.0%; Pred. No. 0.78;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 ENSLDAGAT 38
   |||||||||
Db 38 ENSLDAGAT 46

RESULT 11
D96691
hypothetical protein T12I7.1 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: D96691
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96691
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-249 <STO>
A;Cross-references: UNIPROT:Q9C557; UNIPARC:UPI000009D93C; GB:AE005173; NID:g11054575; P
C;Genetics:
A;Gene: T12I7.1
A;Map position: 1

  Query Match 1.0%; Score 9; DB 2; Length 249;
  Best Local Similarity 100.0%; Pred. No. 1;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 759 KRLLENHKL 767
   |||||||||
Db 24 KRLLENHKL 32

RESULT 12
JC2399
PMS4 homolog mismatch repair protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: JC2399
R;Horii, A.; Han, H.J.; Sasaki, S.; Shimada, M.; Nakamura, Y.
Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994
A;Title: Cloning, characterization and chromosomal assignment of the human genes homolog
A;Reference number: JC2398; MUID:95071462; PMID:7980603
A;Accession: JC2399
A;Molecule type: DNA
A;Residues: 1-252 <HOR>
A;Cross-references: UNIPROT:Q16530; UNIPROT:Q15157; UNIPARC:UPI000017C30A; DDBJ:D38436
C;Genetics:
A;Gene: GDB:PMS2L2; PMS4
A;Cross-references: GDB:437143
A;Map position: 7q11.23-7q22

  Query Match 1.0%; Score 9; DB 2; Length 252;
  Best Local Similarity 100.0%; Pred. No. 1;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 ENSLDAGAT 38
   |||||||||
Db 56 ENSLDAGAT 64

RESULT 13
JC2398
PMS3 homolog mismatch repair protein - human
C;Species: Homo sapiens (man)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C;Accession: JC2398
R;Horii, A.; Han, H.J.; Sasaki, S.; Shimada, M.; Nakamura, Y.
Biochem. Biophys. Res. Commun. 204, 1257-1264, 1994
A;Title: Cloning, characterization and chromosomal assignment of the human genes homol
A;Reference number: JC2398; MUID:95071462; PMID:7980603
A;Accession: JC2398
A;Molecule type: DNA
A;Residues: 1-256 <HOR>
A;Cross-references: UNIPROT:Q16530; UNIPARC:UPI000017C309; DDBJ:D38435
C;Genetics:
A;Gene: GDB:PMS2L1; PMS3
A;Cross-references: GDB:437142
A;Map position: 7q11.23-7q22
C;Keywords: DNA repair

  Query Match 1.0%; Score 9; DB 2; Length 256;
  Best Local Similarity 100.0%; Pred. No. 1;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 30 ENSLDAGAT 38
   |||||||||
Db 60 ENSLDAGAT 68

RESULT 14
D70436
DNA mismatch repair protein MutL - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: D70436
R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320
A;Accession: D70436
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-425 <AQF>
A;Cross-references: UNIPROT:O67518; UNIPARC:UPI0000056665; GB:AE000746; NID:g2983925;
A;Experimental source: strain VF5
C;Genetics:
A;Gene: mutL

  Query Match 1.0%; Score 9; DB 2; Length 425;
  Best Local Similarity 100.0%; Pred. No. 1.6;
  Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 TYGFRGEAL 97
   |||||||||
Db 91 TYGFRGEAL 99

RESULT 15
H72427
DNA mismatch repair protein - Thermotoga maritima (strain MSB8)

C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: H72427
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: H72427
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <ARN>
A;Cross-references: UNIPARC:UPI0000166105; GB:AE001690; GB:AE000512; NID:g4980496; PIDN:
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM0022
C;Superfamily: mismatch repair protein hexB

Query Match 1.0%; Score 9; DB 2; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          89 TYGFRGEAL 97
               |||||||||
Db          97 TYGFRGEAL 105
```

Search completed: December 21, 2005, 20:33:11
Job time : 48 secs

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:21:32 ; Search time 240 Seconds
(without alignments)
2739.802 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 932
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 932 | 100.0 | 932 | 1 | PMS1_HUMAN | P54277 homo sapien |
| 2 | 920 | 98.7 | 920 | 2 | Q4VAL4_HUMAN | Q4val4 homo sapien |
| 3 | 699 | 75.0 | 893 | 2 | Q5FBZ3_HUMAN | Q5fbz3 homo sapien |
| 4 | 618 | 66.3 | 669 | 2 | Q4VAL5_HUMAN | Q4val5 homo sapien |
| 5 | 618 | 66.3 | 770 | 2 | Q5FBZ8_HUMAN | Q5fbz8 homo sapien |
| 6 | 463 | 49.7 | 667 | 2 | Q5FBZ9_HUMAN | Q5fbz9 homo sapien |
| 7 | 385 | 41.3 | 555 | 2 | Q5FBZ6_HUMAN | Q5fbz6 homo sapien |
| 8 | 234 | 25.1 | 234 | 2 | Q5FBZ2_HUMAN | Q5fbz2 homo sapien |
| 9 | 194 | 20.8 | 195 | 2 | Q5FBZ1_HUMAN | Q5fbz1 homo sapien |
| 10 | 194 | 20.8 | 196 | 2 | Q5FBZ5_HUMAN | Q5fbz5 homo sapien |
| 11 | 194 | 20.8 | 248 | 2 | Q5FBZ4_HUMAN | Q5fbz4 homo sapien |
| 12 | 161 | 17.3 | 163 | 2 | Q68DF0_HUMAN | Q68df0 homo sapien |
| 13 | 140 | 15.0 | 165 | 2 | Q5XG96_HUMAN | Q5xg96 homo sapien |
| 14 | 140 | 15.0 | 166 | 2 | Q96HL0_HUMAN | Q96hl0 homo sapien |
| 15 | 137 | 14.7 | 931 | 2 | Q5R904_PONPY | Q5r904 pongo pygma |
| 16 | 44 | 4.7 | 47 | 2 | Q5FBZ7_HUMAN | Q5fbz7 homo sapien |
| 17 | 30 | 3.2 | 917 | 2 | Q8K119_MOUSE | Q8k119 mus musculu |
| 18 | 29 | 3.1 | 919 | 2 | Q6P7D0_RAT | Q6p7d0 rattus norv |
| 19 | 23 | 2.5 | 143 | 2 | Q8BLI9_MOUSE | Q8bli9 mus musculu |
| 20 | 18 | 1.9 | 916 | 2 | Q5ZKT5_CHICK | Q5zkt5 gallus gall |
| 21 | 16 | 1.7 | 580 | 1 | MUTL_CHLPN | Q9z794 chlamydia p |
| 22 | 15 | 1.6 | 342 | 2 | Q6ALT0_DESPS | Q6alt0 desulfotale |
| 23 | 15 | 1.6 | 576 | 1 | MUTL_CHLTR | O84579 chlamydia t |
| 24 | 15 | 1.6 | 1474 | 2 | Q7QNV7_GIALA | Q7qnv7 giardia lam |
| 25 | 15 | 1.6 | 1474 | 2 | Q6WDA1_GIALA | Q6wda1 giardia lam |
| 26 | 14 | 1.5 | 854 | 2 | Q4RTJ3_TETNG | Q4rtj3 tetraodon n |
| 27 | 14 | 1.5 | 925 | 2 | Q7ZXV9_XENLA | Q7zxv9 xenopus lae |
| 28 | 14 | 1.5 | 928 | 2 | Q5FVX9_XENTR | Q5fvx9 xenopus tro |
| 29 | 13 | 1.4 | 194 | 2 | Q8JFW5_BRARE | Q8jfw5 brachydanio |
| 30 | 13 | 1.4 | 372 | 2 | Q7SXD5_BRARE | Q7sxd5 brachydanio |
| 31 | 13 | 1.4 | 724 | 2 | Q5JN46_ORYSA | Q5jn46 oryza sativ |
| 32 | 13 | 1.4 | 831 | 2 | Q4P3V5_USTMA | Q4p3v5 ustilago ma |
| 33 | 13 | 1.4 | 896 | 2 | Q8JFR9_BRARE | Q8jfr9 brachydanio |
| 34 | 13 | 1.4 | 1111 | 2 | Q52FC0_MAGGR | Q52fc0 magnaporthe |
| 35 | 12 | 1.3 | 426 | 1 | MUTL_AQUPY | P70754 aquifex pyr |
| 36 | 12 | 1.3 | 733 | 2 | Q8GY98_ARATH | Q8gy98 arabidopsis |
| 37 | 12 | 1.3 | 779 | 2 | O81287_ARATH | O81287 arabidopsis |
| 38 | 12 | 1.3 | 923 | 2 | Q941I6_ARATH | Q941i6 arabidopsis |
| 39 | 11 | 1.2 | 643 | 2 | Q8RG56_FUSNN | Q8rg56 fusobacteri |
| 40 | 11 | 1.2 | 675 | 2 | Q7P5M3_FUSNV | Q7p5m3 fusobacteri |
| 41 | 11 | 1.2 | 684 | 1 | MLH1_SCHPO | Q9p7w6 schizosacch |
| 42 | 10 | 1.1 | 118 | 2 | Q834Z3_ENTFA | Q834z3 enterococcu |
| 43 | 10 | 1.1 | 574 | 2 | Q83CM9_COXBU | Q83cm9 coxiella bu |
| 44 | 10 | 1.1 | 624 | 1 | MUTL_CHLTE | Q8kax3 chlorobium |
| 45 | 10 | 1.1 | 893 | 2 | O76417_DROME | O76417 drosophila |

ALIGNMENTS

RESULT 1
PMS1_HUMAN
ID PMS1_HUMAN STANDARD; PRT; 932 AA.
AC P54277;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE PMS1 protein homolog 1 (DNA mismatch repair protein PMS1).
GN Name=PMS1; Synonyms=PMSL1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Gall bladder;
RX MEDLINE=94352394; PubMed=8072530; DOI=10.1038/371075a0;
RA Nicolaides N.C., Papadopoulos N., Liu B., Wei Y.-F., Carter K.C.,
RA Ruben S.M., Rosen C.A., Haseltine W.H., Fleischmann R.D., Fraser C.M.,
RA Adams M.D., Venter J.C., Dunlop M.G., Hamilton S.R., Petersen G.M.,
RA de la Chapelle A., Vogelstein B., Kinzler K.W.;
RT "Mutations of two PMS homologues in hereditary nonpolyposis colon
RT cancer.";
RL Nature 371:75-80(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA], AND VARIANTS GLN-27; LYS-202;
RP ARG-501; SER-632; ASP-720 AND HIS-793.
RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RT "NIEHS-SNPs, environmental genome project, NIEHS ES15478, Department
RT of Genome Sciences, Seattle, WA (URL: http://egp.gs.washington.edu).";
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP VARIANTS HNPCC3 THR-394 AND ARG-501.
RX MEDLINE=99408236; PubMed=10480359; DOI=10.1007/s004390051067;
RA Wang Q., Lasset C., Desseigne F., Saurin J.-C., Maugard C.,
RA Navarro C., Ruano E., Descos L., Trillet-Lenoir V., Bosset J.-F.,
RA Puisieux A.;
RT "Prevalence of germline mutations of hMLH1, hMSH2, hPMS1, hPMS2, and
RT hMSH6 genes in 75 French kindreds with nonpolyposis colorectal
RT cancer.";
RL Hum. Genet. 105:79-85(1999).
CC -!- FUNCTION: Probably involved in the repair of mismatches in DNA.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DISEASE: Defects in PMS1 are the cause of hereditary non-polyposis
CC colorectal cancer type 3 (HNPCC3) [MIM:600258]. Mutations in more
CC than one gene locus can be involved alone or in combination in the
CC production of the HNPCC phenotype (also called Lynch syndrome).
CC Most families with clinically recognized HNPCC have mutations in
CC either MLH1 or MSH2 genes. HNPCC is an autosomal, dominantly
CC inherited disease associated with marked increase in cancer
CC susceptibility. It is characterized by a familial predisposition

```
CC       to early onset colorectal carcinoma (CRC) and extra-colonic
CC       cancers of the gastrointestinal, urological and female
CC       reproductive tracts. HNPCC is reported to be the most common form
CC       of inherited colorectal cancer in the Western world, and accounts
CC       for 15% of all colon cancers. Cancers in HNPCC originate within
CC       benign neoplastic polyps termed adenomas. Clinically, HNPCC is
CC       often divided into two subgroups. Type I: hereditary
CC       predisposition to colorectal cancer, a young age of onset, and
CC       carcinoma observed in the proximal colon. Type II: patients have
CC       an increased risk for cancers in certain tissues such as the
CC       uterus, ovary, breast, stomach, small intestine, skin, and larynx
CC       in addition to the colon. Diagnosis of classical HNPCC is based on
CC       the Amsterdam criteria: 3 or more relatives affected by colorectal
CC       cancer, one a first degree relative of the other two; 2 or more
CC       generation affected; 1 or more colorectal cancers presenting
CC       before 50 years of age; exclusion of hereditary polyposis
CC       syndromes. The term "suspected HNPCC" or "incomplete HNPCC" can be
CC       used to describe families who do not or only partially fulfill the
CC       Amsterdam criteria, but in whom a genetic basis for colon cancer
CC       is strongly suspected.
CC   -!- SIMILARITY: Belongs to the DNA mismatch repair mutL/hexB family.
CC   -!- SIMILARITY: Contains 1 HMG box DNA-binding domain.
CC   -!- DATABASE: NAME=Hereditary non-polyposis colorectal cancer db;
CC       WWW="http://www.nfdht.nl/".
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; U13695; AAA63922.1; -; Genomic_DNA.
DR   EMBL; AY267352; AAO89079.1; -; Genomic_DNA.
DR   PIR; S47597; S47597.
DR   HSSP; P54278; 1H7S.
DR   Ensembl; ENSG00000064933; Homo sapiens.
DR   HGNC; HGNC:9121; PMS1.
DR   MIM; 600258; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003677; F:DNA binding; TAS.
DR   GO; GO:0006298; P:mismatch repair; TAS.
DR   InterPro; IPR003594; ATP_bd_ATPase.
DR   InterPro; IPR002099; DNA_mis_repair.
DR   InterPro; IPR000910; HMG_12_box.
DR   PANTHER; PTHR10073; DNA_mis_repair; 1.
DR   Pfam; PF01119; DNA_mis_repair; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00505; HMG_box; 1.
DR   TIGRFAMs; TIGR00585; mutl; 1.
DR   PROSITE; PS00058; DNA_MISMATCH_REPAIR_1; 1.
DR   PROSITE; PS50118; HMG_BOX_2; 1.
KW   Anti-oncogene; Cell cycle; Disease mutation; DNA damage; DNA repair;
KW   Hereditary nonpolyposis colorectal cancer; Nuclear protein;
KW   Polymorphism.
FT   DNA_BIND     571    639       HMG box.
FT   VARIANT       27     27       E -> Q (in dbSNP:5742973).
FT                                 /FTId=VAR_019166.
FT   VARIANT      202    202       R -> K (in dbSNP:2066459).
FT                                 /FTId=VAR_014877.
FT   VARIANT      394    394       M -> T (in incomplete HNPCC3;
FT                                 dbSNP:1145231).
FT                                 /FTId=VAR_012967.
FT   VARIANT      501    501       G -> R (in incomplete HNPCC3;
FT                                 dbSNP:1145232).
FT                                 /FTId=VAR_012968.
FT   VARIANT      632    632       N -> S (in dbSNP:2066456).
FT                                 /FTId=VAR_014878.
FT   VARIANT      720    720       E -> D (in dbSNP:2066455).
FT                                 /FTId=VAR_014879.
FT   VARIANT      793    793       Y -> H (in dbSNP:1145234).
FT                                 /FTId=VAR_014880.
SQ   SEQUENCE   932 AA;  105830 MW;  EC4F402937B616DF CRC64;
```

```
  Query Match           100.0%;  Score 932;  DB 1;  Length 932;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  932;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932


RESULT 2
Q4VAL4_HUMAN
ID   Q4VAL4_HUMAN PRELIMINARY;      PRT;   920 AA.
AC   Q4VAL4;
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:03:01 ; Search time 47 Seconds
(without alignments)
1639.441 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 2 | US-08-294-312B-4 | Sequence 4, Appli |
| 2 | 4812 | 100.0 | 932 | 2 | US-08-468-024B-4 | Sequence 4, Appli |
| 3 | 4812 | 100.0 | 932 | 2 | US-09-708-200-11 | Sequence 11, Appl |
| 4 | 4812 | 100.0 | 932 | 2 | US-08-465-679-4 | Sequence 4, Appli |
| 5 | 4812 | 100.0 | 932 | 2 | US-09-788-657-17 | Sequence 17, Appl |
| 6 | 4812 | 100.0 | 932 | 2 | US-09-788-657-18 | Sequence 18, Appl |
| 7 | 4812 | 100.0 | 932 | 2 | US-09-712-691-9 | Sequence 9, Appli |
| 8 | 4812 | 100.0 | 932 | 2 | US-10-641-068-17 | Sequence 17, Appl |
| 9 | 4812 | 100.0 | 932 | 2 | US-10-641-068-18 | Sequence 18, Appl |
| 10 | 4805 | 99.9 | 932 | 2 | US-09-707-468C-9 | Sequence 9, Appli |
| 11 | 4693 | 97.5 | 921 | 2 | US-09-949-016-7787 | Sequence 7787, Ap |
| 12 | 661 | 13.7 | 133 | 2 | US-09-788-657-21 | Sequence 21, Appl |
| 13 | 661 | 13.7 | 133 | 2 | US-09-712-691-15 | Sequence 15, Appl |
| 14 | 661 | 13.7 | 133 | 2 | US-10-641-068-21 | Sequence 21, Appl |
| 15 | 572 | 11.9 | 864 | 1 | US-08-209-521-29 | Sequence 29, Appl |
| 16 | 572 | 11.9 | 864 | 2 | US-09-265-503B-138 | Sequence 138, App |
| 17 | 569 | 11.8 | 856 | 2 | US-08-709-784-2 | Sequence 2, Appli |
| 18 | 569 | 11.8 | 862 | 1 | US-08-209-521-23 | Sequence 23, Appl |
| 19 | 569 | 11.8 | 862 | 1 | US-08-209-521-30 | Sequence 30, Appl |
| 20 | 569 | 11.8 | 862 | 2 | US-09-059-461-2 | Sequence 2, Appli |
| 21 | 569 | 11.8 | 862 | 2 | US-08-961-810-133 | Sequence 133, App |
| 22 | 569 | 11.8 | 862 | 2 | US-08-352-902D-133 | Sequence 133, App |
| 23 | 569 | 11.8 | 862 | 2 | US-08-294-312B-6 | Sequence 6, Appli |
| 24 | 569 | 11.8 | 862 | 2 | US-08-468-024B-6 | Sequence 6, Appli |
| 25 | 569 | 11.8 | 862 | 2 | US-09-265-503B-133 | Sequence 133, App |
| 26 | 569 | 11.8 | 862 | 2 | US-09-708-200-9 | Sequence 9, Appli |
| 27 | 569 | 11.8 | 862 | 2 | US-08-465-679-6 | Sequence 6, Appli |
| 28 | 569 | 11.8 | 862 | 2 | US-09-712-691-7 | Sequence 7, Appli |
| 29 | 569 | 11.8 | 862 | 2 | US-09-707-468C-7 | Sequence 7, Appli |
| 30 | 569 | 11.8 | 862 | 2 | US-09-749-601A-11 | Sequence 11, Appl |
| 31 | 543.5 | 11.3 | 903 | 1 | US-08-209-521-24 | Sequence 24, Appl |
| 32 | 543.5 | 11.3 | 903 | 2 | US-08-961-810-134 | Sequence 134, App |
| 33 | 543.5 | 11.3 | 903 | 2 | US-08-352-902D-134 | Sequence 134, App |
| 34 | 543.5 | 11.3 | 903 | 2 | US-09-265-503B-134 | Sequence 134, App |
| 35 | 536 | 11.1 | 779 | 2 | US-09-749-601A-12 | Sequence 12, Appl |
| 36 | 534.5 | 11.1 | 859 | 2 | US-09-708-200-7 | Sequence 7, Appli |
| 37 | 534.5 | 11.1 | 859 | 2 | US-09-788-657-16 | Sequence 16, Appl |
| 38 | 534.5 | 11.1 | 859 | 2 | US-09-712-691-5 | Sequence 5, Appli |
| 39 | 534.5 | 11.1 | 859 | 2 | US-09-707-468C-5 | Sequence 5, Appli |
| 40 | 534.5 | 11.1 | 859 | 2 | US-10-641-068-16 | Sequence 16, Appl |
| 41 | 445.5 | 9.3 | 607 | 1 | US-08-209-521-7 | Sequence 7, Appli |
| 42 | 445.5 | 9.3 | 607 | 2 | US-08-961-810-3 | Sequence 3, Appli |
| 43 | 445.5 | 9.3 | 607 | 2 | US-08-352-902D-3 | Sequence 3, Appli |
| 44 | 445.5 | 9.3 | 607 | 2 | US-09-265-503B-3 | Sequence 3, Appli |
| 45 | 425.5 | 8.8 | 769 | 2 | US-09-788-657-15 | Sequence 15, Appl |

ALIGNMENTS

RESULT 1
US-08-294-312B-4
; Sequence 4, Application US/08294312B
; Patent No. 6380369
; GENERAL INFORMATION:
; APPLICANT: Adams et al.
; TITLE OF INVENTION: Human DNA Mismatch Repair Proteins
; FILE REFERENCE: PF106P2
; CURRENT APPLICATION NUMBER: US/08/294,312B
; CURRENT FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/210,143
; PRIOR FILING DATE: 1994-03-16
; PRIOR APPLICATION NUMBER: 08/187,757
; PRIOR FILING DATE: 1994-01-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 932
; TYPE: PRT
; ORGANISM: homo sapiens
US-08-294-312B-4

Query Match 100.0%; Score 4812; DB 2; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
```

```
Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
US-08-468-024B-4
; Sequence 4, Application US/08468024B
; Patent No. 6416984
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Human DNA Mismatch Repair Proteins
; FILE REFERENCE: PF106P3
; CURRENT APPLICATION NUMBER: US/08/468,024B
; CURRENT FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/294,312
; PRIOR FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/210,143
; PRIOR FILING DATE: 1994-03-16
; PRIOR APPLICATION NUMBER: 08/187,757
; PRIOR FILING DATE: 1994-01-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 932
; TYPE: PRT
; ORGANISM: homo sapiens
US-08-468-024B-4

Query Match 100.0%; Score 4812; DB 2; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 3
US-09-708-200-11
; Sequence 11, Application US/09708200
; Patent No. 6576468
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas C
; APPLICANT: Grasso, Luigi
; APPLICANT: Sass, Philip M
; TITLE OF INVENTION: METHODS FOR ISOLATING NOVEL ANTIMICROBIAL AGENTS FROM
; TITLE OF INVENTION: HYPERMUTABLE CELLS
; FILE REFERENCE: MOR-0005
; CURRENT APPLICATION NUMBER: US/09/708,200

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 19:59:55 ; Search time 46 Seconds
(without alignments)
1949.435 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 2 | S47597 | mutL protein homol |
| 2 | 587.5 | 12.2 | 805 | 2 | T21957 | hypothetical prote |
| 3 | 569 | 11.8 | 862 | 2 | S47598 | mutL protein homol |
| 4 | 556 | 11.6 | 904 | 2 | S53896 | DNA mismatch repai |
| 5 | 536 | 11.1 | 779 | 2 | T01304 | hypothetical prote |
| 6 | 529 | 11.0 | 794 | 2 | T37989 | DNA mismatch repai |
| 7 | 448 | 9.3 | 669 | 2 | C89904 | DNA mismatch repai |
| 8 | 424.5 | 8.8 | 769 | 2 | S54525 | mismatch repair pr |
| 9 | 403 | 8.4 | 615 | 2 | PH0853 | methyl-directed mi |
| 10 | 403 | 8.4 | 615 | 2 | B91272 | enzyme in methyl-d |
| 11 | 403 | 8.4 | 615 | 2 | B86113 | enzyme in methyl-d |
| 12 | 401.5 | 8.3 | 635 | 2 | AC0046 | DNA mismatch repai |
| 13 | 396.5 | 8.2 | 516 | 2 | H72427 | DNA mismatch repai |
| 14 | 394.5 | 8.2 | 629 | 2 | E64046 | mismatch repair pr |
| 15 | 394 | 8.2 | 779 | 2 | T25389 | hypothetical prote |
| 16 | 392 | 8.1 | 610 | 2 | C70126 | DNA mismatch repai |
| 17 | 392 | 8.1 | 618 | 2 | AG1048 | DNA mismatch repai |
| 18 | 392 | 8.1 | 618 | 2 | A33588 | mismatch repair pr |
| 19 | 390.5 | 8.1 | 684 | 2 | T50317 | probable DNA misma |
| 20 | 389.5 | 8.1 | 601 | 2 | AD1250 | DNA mismatch repai |
| 21 | 387.5 | 8.1 | 603 | 2 | AH1612 | DNA mismatch repai |
| 22 | 387.5 | 8.1 | 653 | 2 | A82334 | DNA mismatch repai |
| 23 | 385 | 8.0 | 619 | 2 | E82765 | DNA mismatch repai |
| 24 | 385 | 8.0 | 756 | 2 | S43085 | DNA mismatch repai |
| 25 | 377.5 | 7.8 | 637 | 2 | H83945 | DNA mismatch repai |
| 26 | 368 | 7.6 | 627 | 2 | A69663 | DNA mismatch repai |
| 27 | 367 | 7.6 | 622 | 2 | E97126 | DNA mismatch repai |
| 28 | 367 | 7.6 | 695 | 2 | H86900 | DNA mismatch repai |
| 29 | 363.5 | 7.6 | 425 | 2 | D70436 | DNA mismatch repai |
| 30 | 363.5 | 7.6 | 595 | 2 | F71650 | DNA mismatch repai |
| 31 | 361.5 | 7.5 | 584 | 2 | D84996 | DNA mismatch repai |
| 32 | 359.5 | 7.5 | 633 | 2 | F83028 | DNA mismatch repai |
| 33 | 356 | 7.4 | 737 | 2 | F85092 | MLH1 protein [impo |
| 34 | 356 | 7.4 | 737 | 2 | T51620 | DNA mismatch repai |
| 35 | 355.5 | 7.4 | 576 | 2 | A71497 | probable DNA misma |
| 36 | 355 | 7.4 | 649 | 2 | H97891 | DNA mismatch repai |
| 37 | 355 | 7.4 | 649 | 2 | A95020 | DNA mismatch repai |
| 38 | 352.5 | 7.3 | 658 | 2 | C81860 | DNA mismatch repai |
| 39 | 350 | 7.3 | 649 | 2 | A33589 | mismatch repair pr |
| 40 | 349.5 | 7.3 | 658 | 2 | B81084 | mismatch repair pr |
| 41 | 344.5 | 7.2 | 580 | 2 | A72032 | DNA mismatch repai |
| 42 | 344.5 | 7.2 | 580 | 2 | B86592 | DNA mismatch repai |
| 43 | 344 | 7.1 | 576 | 2 | G81657 | DNA mismatch repai |
| 44 | 340 | 7.1 | 695 | 2 | S64862 | hypothetical prote |
| 45 | 332 | 6.9 | 610 | 2 | A97870 | DNA mismatch repai |

ALIGNMENTS

RESULT 1
S47597
mutL protein homolog - human
C;Species: Homo sapiens (man)
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C;Accession: S47597
R;Nicolaides, N.C.; Papadopoulos, N.; Liu, B.; Wei, Y.F.; Carter, K.C.; Ruben, S.M.; R
S.R.; Petersen, G.M.; de la Chapelle, A.; Vogelstein, B.; Kinzler, K.W.
Nature 371, 75-80, 1994
A;Title: Mutations of two PMS homologues in hereditary nonpolyposis colon cancer.
A;Reference number: S47597; MUID:94352394; PMID:8072530
A;Accession: S47597
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-932 <NIC>
A;Cross-references: UNIPROT:P54277; UNIPARC:UPI00000405F5; EMBL:U13695; NID:g535512; P
C;Genetics:
A;Gene: GDB:PMS1; PMSL1
A;Cross-references: GDB:386403; OMIM:600258
A;Map position: 2q31-2q33
F;571-643/Domain: HMG box homology <HMG1>

Query Match 100.0%; Score 4812; DB 2; Length 932;
Best Local Similarity 100.0%; Pred. No. 7.3e-231;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
T21957
hypothetical protein H12C20.2a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T21957; T23069
R;Lennard, N.
submitted to the EMBL Data Library, June 1996
A;Reference number: Z19493
A;Accession: T21957
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-805 <WIL>
A;Cross-references: UNIPROT:Q9TVL8; UNIPARC:UPI0000081DBE; EMBL:Z74033; PIDN:CAA98478.1;
A;Experimental source: clone F38B7
R;White, S.
submitted to the EMBL Data Library, March 1998
A;Reference number: Z19670
A;Accession: T23069
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-805 <WI2>
A;Cross-references: UNIPARC:UPI0000081DBE; EMBL:AL022272; PIDN:CAA18355.1; GSPDB:GN00023
A;Experimental source: clone H12C20
C;Genetics:
A;Gene: CESP:H12C20.2a
A;Map position: 5
A;Introns: 21/3; 109/2; 267/2; 300/3; 329/1; 393/2; 553/3; 612/2; 668/2; 724/1; 783/3
C;Superfamily: DNA mismatch repair protein

  Query Match            12.2%;  Score 587.5;  DB 2;  Length 805;
  Best Local Similarity  22.9%;  Pred. No. 1.1e-21;
  Matches  226;  Conservative  164;  Mismatches  326;  Indels  271;  Gaps  37;

```
Qy         12 LSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEGIKAVDAPVMAM 71
              |:::|:: |: | :::||:||:|||:| :|::::| ||: |||:||| ||:| :   :
Db         17 LTTAQVVVSLSSAIRQLIDNSIDAGSTIIDIRVKNNGFESIEVQDNGSGIEARNFDALCK 76

Qy         72 KYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQYVLDGSGHILS 131
              : |||:   | : | | ||||||| ::| :: | | || :     |:   | ||:|:
Db         77 PHSTSKLTQFSDFDKLATLGFRGEALNALCTVSSVSIFTRASDTEIGTRLTYDHSGNIIC 136

Qy        132 QKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFGILKPDLRIVFV 191
              :: :    |||:   :||: ||||::     :  | |  |:   : || :| | ::|:
Db        137 RQSAARELGTTIIVNKLFETLPVRRK--ELERSQKREFVKLLSTVQSFALLCPHIKILCT 194

Qy        192 HNKAVIWQKSRVSDHKMALMSVLG--TAVMN---------------------------- 220
              :|          ::  |  |:    |  |:: :
Db        195 NN---------INGKKTNLICTPGGTTSIQDVVANLFGIARKIENSKIGSGLIPIQQNQP 245

Qy        221 NMESFQYHS---EESQIY----LSGFLPKCDADHSFTSLSTPERSFIFINSRPVHQKDIL 273
              ::|    ||   ||   :    : ||:  |  :|       | :|  |::||:|||    :
Db        246 DVEIMTIHSVPMEEMHFFDLFKIRGFVSSC--EHG-CGRGTSDRQFVYINNRPVEYSRVC 302

Qy        274 KLIRHHYNLKCLKESTRLYPVFFLKIDVPTADVDVNLTPDKSQVLLQNKESVLIALENLM 333
              :|    |     : : : ||:  | ||||   :|||:||||  |:|: :  :|  :   |
Db        303 SVINDVYK----QFNKKQYPIIVLFIDVPPEKIDVNVTPDKKTVMLEKERHLLAVVRASM 358

Qy        334 TTCYGPLPSTNSYENNKTDVSAADIV-LSKTAETDVLFNKVESSGKNYSNVDTSVIPFQN 392
              |  |    |:   :: |    |: || : :: |   :|| :  | |       |
Db        359 MKTY--LKIVGSHSTVRSSVEDRRIMNLSQQSFSNASFMSSKSSTPDDFNNTT-----LN 411

Qy        393 DMHNDESGKNTDDCLNHQISIGDFGYGHCSSEISNIDKNTKNAFQDISMSNVSWENSQTE 452
              : ::|  || | |                    | |             ||
Db        412 STYPEDSLLNTSDLL----------------------KQRK------------ENRSPP 436

Qy        453 YSKTCFISSVKHTQSENGNKDHIDESGENEEEAGLENSS--------EIS---------- 494
              |:|    :: |:          |   |  |     ||| |         |:|
Db        437 AKKSC--PMIRRTE-----PFHSVPSTSNSRTQRLENFSFTMEPKRVEVSKKIPSKSDKK 489

Qy        495 -ADEWSRGNILKNS----VGENIEPVKILVPEKSLPCKVSNNNYPIPEQMNLNEDSCNKK 549
               ||  |  ::: :     || |: ::||  |:|              |  ::|| |::
Db        490 LTDEELRSAVIEENPLKKAGE-IDDIEIL--EQS------------QESQDVNESQCSQD 534

Qy        550 SNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLEDATLQIEELW 609
              |    |  :|: : |  : ||  |    |  : : |:  :    : |:| :  |
Db        535 SQTSQN--SRVSYFTLRPQQKIKFSMK---LLREAYSPK--TDETDDNTEEAEVSAE--- 584

Qy        610 KTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAWNLAQKHKLK 669
              | :  |   |  ::   | ||   |: |::          ||   |:|    :|
Db        585 KDVLNEITTKINKEENDDAER---QLSRSL-------TKDDFSKMKIIGQFN-------- 626

Qy        670 TSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEEKDEPCLIHN 729
                                                                  :|
Db        627 ----------------------------------------------------HGFIICR 633

Qy        730 LRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTESLFNGSHYL 789
              ||          : ::: :  :|   |:||  : ||  :||  |  |     |     |
Db        634 LR----------GHLFIVDQHASDEKYNFERLQSSAKLTKQPLFMPTALG---FGAVQEL 680

Qy        790 DVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEI------EGMANCLP 843
               :  :                |  |||| :          : :|         : |
Db        681 IIRENL--------------PIFHANGFDFEFSENDGCIKTFLTARPELLNQQLTN--- 722

Qy        844 FYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDIIYR 903
              :||:||| |::::   ::|  || ::      :| | |  :  |:: ::  ||
Db        723 ----SDLEEIL-AVVSQYPNQMY--RPVRIRKIFASKACRKSVMIGKPLNQREMTQIIRH 775

Qy        904 MKHQFGNEIKECVHGRPFFHHLTYLPE 930
              :    ::   | ||||   ||  ||:
Db        776 LAKL--DQPWNCPHGRPTIRHLASLPD 800
```

RESULT 3
S47598
mutL protein homolog - human
C;Species: Homo sapiens (man)

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 19:58:15 ; Search time 190 Seconds
(without alignments)
2155.269 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 4 | AAB85852 | Aab85852 Human PMS |
| 2 | 4812 | 100.0 | 932 | 4 | AAG63953 | Aag63953 Amino aci |
| 3 | 4812 | 100.0 | 932 | 4 | AAG63954 | Aag63954 Amino aci |
| 4 | 4812 | 100.0 | 932 | 5 | AAU98776 | Aau98776 Human pos |
| 5 | 4812 | 100.0 | 932 | 5 | AAE24357 | Aae24357 Human mis |
| 6 | 4812 | 100.0 | 932 | 5 | AAO18553 | Aao18553 Human mis |
| 7 | 4812 | 100.0 | 932 | 5 | AAE28277 | Aae28277 Human MLH |
| 8 | 4812 | 100.0 | 932 | 5 | AAE24684 | Aae24684 Human PMS |
| 9 | 4812 | 100.0 | 932 | 6 | ABU07972 | Abu07972 Human PMS |
| 10 | 4812 | 100.0 | 932 | 6 | ABU07971 | Abu07971 Human PMS |
| 11 | 4812 | 100.0 | 932 | 6 | ABU89659 | Abu89659 Human PMS |
| 12 | 4812 | 100.0 | 932 | 6 | ABU89660 | Abu89660 Human PMS |
| 13 | 4812 | 100.0 | 932 | 6 | ABO07415 | Abo07415 Human Mut |
| 14 | 4812 | 100.0 | 932 | 6 | AAO27514 | Aao27514 Human mis |
| 15 | 4812 | 100.0 | 932 | 6 | AAO27515 | Aao27515 Human mis |
| 16 | 4812 | 100.0 | 932 | 7 | ADA06246 | Ada06246 Human mis |
| 17 | 4812 | 100.0 | 932 | 7 | ADA06244 | Ada06244 Human mis |
| 18 | 4812 | 100.0 | 932 | 7 | ADC89605 | Adc89605 Human PMS |
| 19 | 4812 | 100.0 | 932 | 7 | ADC89607 | Adc89607 Human PMS |
| 20 | 4812 | 100.0 | 932 | 7 | ADF17892 | Adf17892 Human PMS |
| 21 | 4812 | 100.0 | 932 | 7 | ADG62881 | Adg62881 Human PMS |
| 22 | 4812 | 100.0 | 932 | 7 | ADG62882 | Adg62882 Human PMS |
| 23 | 4812 | 100.0 | 932 | 7 | ADH62629 | Adh62629 Human mis |
| 24 | 4812 | 100.0 | 932 | 7 | ADH60983 | Adh60983 Human mis |
| 25 | 4812 | 100.0 | 932 | 7 | ADJ68675 | Adj68675 Human hea |
| 26 | 4812 | 100.0 | 932 | 8 | ADF78840 | Adf78840 Human mis |
| 27 | 4812 | 100.0 | 932 | 8 | ADG46767 | Adg46767 Human MMR |
| 28 | 4812 | 100.0 | 932 | 8 | ADO40066 | Ado40066 Human PMS |
| 29 | 4812 | 100.0 | 932 | 8 | ADP66682 | Adp66682 Human mis |
| 30 | 4812 | 100.0 | 932 | 8 | ADP66680 | Adp66680 Human mis |
| 31 | 4812 | 100.0 | 932 | 8 | ADR13883 | Adr13883 Human DNA |
| 32 | 4812 | 100.0 | 932 | 8 | ADU77038 | Adu77038 Human mis |
| 33 | 4812 | 100.0 | 932 | 9 | ADX08123 | Adx08123 Cyclin-de |
| 34 | 4812 | 100.0 | 932 | 9 | ADX58473 | Adx58473 Amino aci |
| 35 | 4812 | 100.0 | 932 | 9 | ADY53431 | Ady53431 Human PMS |
| 36 | 4807 | 99.9 | 931 | 2 | AAR79009 | Aar79009 Human DNA |
| 37 | 4805 | 99.9 | 932 | 8 | ADT98686 | Adt98686 Human PMS |
| 38 | 4805 | 99.9 | 932 | 8 | ADU77022 | Adu77022 Human mis |
| 39 | 4627 | 96.2 | 900 | 4 | AAB85851 | Aab85851 Human PMS |
| 40 | 4317.5 | 89.7 | 847 | 8 | ABM83683 | Abm83683 Human dia |
| 41 | 1168 | 24.3 | 232 | 5 | AAE28280 | Aae28280 Human MLH |
| 42 | 661 | 13.7 | 133 | 4 | AAB85855 | Aab85855 Human PMS |
| 43 | 661 | 13.7 | 133 | 4 | AAG63957 | Aag63957 Amino aci |
| 44 | 661 | 13.7 | 133 | 5 | AAO18559 | Aao18559 Human tru |
| 45 | 661 | 13.7 | 133 | 5 | AAE24687 | Aae24687 Human PMS |

ALIGNMENTS

RESULT 1
AAB85852
ID AAB85852 standard; protein; 932 AA.
XX
AC AAB85852;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PMS1 protein.
XX
KW Hypermutable bacteria; mismatch repair gene; MMR gene; MutH; MutS; MutL;
KW MutY; PMS2; MLH1; MLH3; PMSR; biocatalysis; bioremediation; biochemical;
KW drug discovery; detoxification; toxin; biotransformation; PMS1.
XX
OS Homo sapiens.
XX
PN WO200159092-A2.
XX
PD 16-AUG-2001.
XX
PF 12-FEB-2001; 2001WO-US004339.
XX
PR 11-FEB-2000; 2000US-0181929P.
XX
PA (UYJO ) UNIV JOHNS HOPKINS.
XX
PI Nicolaides NC, Sass PM, Grasso L, Vogelstein B, Kinzler KW;
XX
DR WPI; 2001-514664/56.
DR N-PSDB; AAH76365.
XX
PT Making hypermutable bacteria for biocatalysis, bioremediation and drug
PT discovery, involves introducing polynucleotide comprising dominant
PT negative allele of mismatch repair gene under regulatory sequence
PT control.
XX
PS Example 1; Page 41; 68pp; English.
XX
CC The invention provides a method for generating a hypermutable bacteria.
CC The method involves introducing a polynucleotide having a dominant
CC negative allele of a mismatch repair (MMR) gene under the control of an
CC inducible transcription regulatory sequence, into a bacterium. The cell
CC becomes inducibly hypermutable. The method is useful to create desirable
CC output traits for commercial applications, using dominant negative
CC alleles of mismatch repair proteins. The mismatch repair gene is a MutH,
CC MutS, MutL or MutY homologue and can be selected from PMS2, MLH1, MLH3,
CC PMSR or PMSR homologue. The hypermutable bacteria is useful for the

CC production, biocatalysis, bioremediation and drug discovery. It is also
CC useful in manufacturing industry for the generation of new biochemicals
CC useful for detoxifying noxious chemicals from by-products of
CC manufacturing processes or those used as catalysts, for remediation of
CC toxins present in the environment including polychlorobenzenes, heavy
CC metals and other environmental hazards for which there is a need to
CC remove them from the environment. The hypermutable bacteria is also
CC useful for screening novel mutations in a gene or a set of genes that
CC produce variant siblings that exhibit a new output trait not found in
CC wild type cells. The bacteria are also useful for producing increased
CC quantity or quality of protein or non-protein therapeutic molecule e.g.
CC Penicillin G, Erythromycin and Clavulanic acid, by biotransformation.
CC Dominant negative alleles of the MMR gene are useful for producing higher
CC quantities of recombinant polypeptides. The present sequence represents a
CC human PMS1 protein
XX
SQ Sequence 932 AA;

Query Match 100.0%; Score 4812; DB 4; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
AAG63953
ID AAG63953 standard; protein; 932 AA.
XX
AC AAG63953;
XX
DT 29-OCT-2001 (first entry)
XX
DE Amino acid sequence of human mismatch repair protein PMS2.
XX
KW PMS2; mismatch repair gene; MMR gene; hypermutable yeast.
XX
OS Homo sapiens.
XX
PN WO200162945-A1.
XX
PD 30-AUG-2001.
XX
PF 21-FEB-2001; 2001WO-US005447.
XX
PR 23-FEB-2000; 2000US-0184336P.
XX
PA (UYJO ) UNIV JOHNS HOPKINS.
PA (NICO/) NICOLAIDES N C.
PA (SASS/) SASS P M.
PA (GRAS/) GRASSO L.
PA (VOGE/) VOGELSTEIN B.
PA (KINZ/) KINZLER K W.
XX
PI Nicolaides NC, Sass PM, Grasso L, Vogelstein B, Kinzler KW;
XX
DR WPI; 2001-522820/57.
DR N-PSDB; AAH75041.
XX
PT Making hypermutable yeast that exhibit novel selected output traits for
PT commercial applications, comprises introducing polynucleotide containing
PT dominant negative allele of mismatch repair gene.
XX
PS Disclosure; Page 38; 60pp; English.
XX
CC The present sequence represents human PSM2. PMS2 is a mismatch repair
CC (MMR) gene. The specification describes a method for making a
CC hypermutable yeast, comprising introducing a polynucleotide containing a
CC dominant negative allele of a mismatch repair (MMR) gene, into a yeast,
CC whereby the cell becomes hypermutable. The method is useful to create
CC desirable output traits for commercial applications, using dominant
CC negative alleles of mismatch repair proteins. The hypermutable yeast is
CC useful for production, biocatalysis, bioremediation and drug discovery.
CC It is also useful in genetic screens for the direct selection of variant
CC subclones that exhibit new output traits. The hypermutable yeast is also
CC useful in the manufacturing industry for the generation of new
CC biochemicals, for detoxifying noxious chemicals from by-products of
CC manufacturing processes or those used as catalysts, for remediation of
CC toxins present in the environment including polychlorobenzenes, heavy
CC metals and other environmental hazards for which there is a need to
CC remove them from the environment. The yeast is also useful for producing
CC increased quantity or quality of protein or non-protein therapeutic
CC molecule e.g., Penicillin G, Erythromycin and Clavulanic acid, by
CC biotransformation

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:06:26 ; Search time 167 Seconds
(without alignments)
2331.838 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 3 | US-09-788-657-17 | Sequence 17, Appl |
| 2 | 4812 | 100.0 | 932 | 3 | US-09-788-657-18 | Sequence 18, Appl |
| 3 | 4812 | 100.0 | 932 | 3 | US-09-912-697-6 | Sequence 6, Appli |
| 4 | 4812 | 100.0 | 932 | 3 | US-09-912-697-8 | Sequence 8, Appli |
| 5 | 4812 | 100.0 | 932 | 3 | US-09-760-285-18 | Sequence 18, Appl |
| 6 | 4812 | 100.0 | 932 | 4 | US-10-079-429-4 | Sequence 4, Appli |
| 7 | 4812 | 100.0 | 932 | 4 | US-10-270-839-27 | Sequence 27, Appl |
| 8 | 4812 | 100.0 | 932 | 4 | US-10-270-839-29 | Sequence 29, Appl |
| 9 | 4812 | 100.0 | 932 | 4 | US-10-243-130-9 | Sequence 9, Appli |
| 10 | 4812 | 100.0 | 932 | 4 | US-10-371-857-3 | Sequence 3, Appli |
| 11 | 4812 | 100.0 | 932 | 4 | US-10-371-857-4 | Sequence 4, Appli |
| 12 | 4812 | 100.0 | 932 | 4 | US-10-371-634-5 | Sequence 5, Appli |
| 13 | 4812 | 100.0 | 932 | 4 | US-10-371-634-7 | Sequence 7, Appli |
| 14 | 4812 | 100.0 | 932 | 4 | US-10-348-074-3 | Sequence 3, Appli |
| 15 | 4812 | 100.0 | 932 | 4 | US-10-348-074-5 | Sequence 5, Appli |
| 16 | 4812 | 100.0 | 932 | 4 | US-10-369-845-11 | Sequence 11, Appl |
| 17 | 4812 | 100.0 | 932 | 4 | US-10-641-068-17 | Sequence 17, Appl |
| 18 | 4812 | 100.0 | 932 | 4 | US-10-641-068-18 | Sequence 18, Appl |
| 19 | 4812 | 100.0 | 932 | 4 | US-10-408-765A-481 | Sequence 481, App |
| 20 | 4812 | 100.0 | 932 | 4 | US-10-813-502-9 | Sequence 9, Appli |
| 21 | 4812 | 100.0 | 932 | 4 | US-10-714-228-2 | Sequence 2, Appli |
| 22 | 4812 | 100.0 | 932 | 4 | US-10-714-228-4 | Sequence 4, Appli |
| 23 | 4812 | 100.0 | 932 | 5 | US-10-933-034-2 | Sequence 2, Appli |
| 24 | 4812 | 100.0 | 932 | 5 | US-10-901-650-9 | Sequence 9, Appli |
| 25 | 4812 | 100.0 | 932 | 6 | US-11-056-776-13 | Sequence 13, Appl |
| 26 | 4805 | 99.9 | 932 | 5 | US-10-850-370-9 | Sequence 9, Appli |
| 27 | 661 | 13.7 | 133 | 3 | US-09-788-657-21 | Sequence 21, Appl |
| 28 | 661 | 13.7 | 133 | 3 | US-09-912-697-14 | Sequence 14, Appl |
| 29 | 661 | 13.7 | 133 | 3 | US-09-760-285-24 | Sequence 24, Appl |
| 30 | 661 | 13.7 | 133 | 4 | US-10-270-839-35 | Sequence 35, Appl |
| 31 | 661 | 13.7 | 133 | 4 | US-10-371-857-12 | Sequence 12, Appl |
| 32 | 661 | 13.7 | 133 | 4 | US-10-371-634-13 | Sequence 13, Appl |
| 33 | 661 | 13.7 | 133 | 4 | US-10-348-074-11 | Sequence 11, Appl |
| 34 | 661 | 13.7 | 133 | 4 | US-10-641-068-21 | Sequence 21, Appl |
| 35 | 661 | 13.7 | 133 | 4 | US-10-714-228-6 | Sequence 6, Appli |
| 36 | 615.5 | 12.8 | 899 | 6 | US-11-097-143-15705 | Sequence 15705, A |
| 37 | 607 | 12.6 | 923 | 4 | US-10-270-839-45 | Sequence 45, Appl |
| 38 | 607 | 12.6 | 923 | 4 | US-10-714-228-36 | Sequence 36, Appl |
| 39 | 607 | 12.6 | 923 | 4 | US-10-714-228-40 | Sequence 40, Appl |
| 40 | 607 | 12.6 | 923 | 5 | US-10-933-034-36 | Sequence 36, Appl |
| 41 | 587.5 | 12.2 | 805 | 4 | US-10-369-493-6498 | Sequence 6498, Ap |
| 42 | 587.5 | 12.2 | 805 | 4 | US-10-369-493-6499 | Sequence 6499, Ap |
| 43 | 585.5 | 12.2 | 1013 | 4 | US-10-425-114-62661 | Sequence 62661, A |
| 44 | 582.5 | 12.1 | 990 | 4 | US-10-425-115-362659 | Sequence 362659, |
| 45 | 572 | 11.9 | 864 | 4 | US-10-349-607-138 | Sequence 138, App |

ALIGNMENTS

RESULT 1
US-09-788-657-17
; Sequence 17, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-657-17

Query Match 100.0%; Score 4812; DB 3; Length 932;
Best Local Similarity 100.0%; Pred. No. 1.5e-285;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
```

```
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
US-09-788-657-18
; Sequence 18, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-657-18

Query Match 100.0%; Score 4812; DB 3; Length 932;
Best Local Similarity 100.0%; Pred. No. 1.5e-285;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 3
US-09-912-697-6
; Sequence 6, Application US/09912697
; Publication No. US20030068808A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas C
; APPLICANT: Sass, Philip M
; APPLICANT: Grasso, Luigi M

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:11:22 ; Search time 13 Seconds
(without alignments)
511.317 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 4812
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 53982 seqs, 7132107 residues

Total number of hits satisfying chosen parameters: 53982

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 4812 | 100.0 | 932 | 7 | US-11-188-743-17 | Sequence 17, Appl |
| 2 | 4812 | 100.0 | 932 | 7 | US-11-188-743-18 | Sequence 18, Appl |
| 3 | 661 | 13.7 | 133 | 7 | US-11-188-743-21 | Sequence 21, Appl |
| 4 | 569 | 11.8 | 862 | 7 | US-11-128-420-11 | Sequence 11, Appl |
| 5 | 536 | 11.1 | 779 | 7 | US-11-128-420-12 | Sequence 12, Appl |
| 6 | 534.5 | 11.1 | 859 | 7 | US-11-188-743-16 | Sequence 16, Appl |
| 7 | 425.5 | 8.8 | 769 | 7 | US-11-188-743-15 | Sequence 15, Appl |
| 8 | 385 | 8.0 | 756 | 7 | US-11-188-743-20 | Sequence 20, Appl |
| 9 | 356 | 7.4 | 737 | 7 | US-11-128-420-9 | Sequence 9, Appli |
| 10 | 355.5 | 7.4 | 658 | 6 | US-10-467-657-4782 | Sequence 4782, Ap |
| 11 | 309.5 | 6.4 | 389 | 7 | US-11-188-743-23 | Sequence 23, Appl |
| 12 | 274 | 5.7 | 133 | 7 | US-11-128-420-13 | Sequence 13, Appl |
| 13 | 265.5 | 5.5 | 147 | 7 | US-11-128-420-14 | Sequence 14, Appl |
| 14 | 249 | 5.2 | 1151 | 7 | US-11-128-420-10 | Sequence 10, Appl |
| 15 | 168.5 | 3.5 | 1404 | 6 | US-10-878-556A-169 | Sequence 169, App |
| 16 | 165.5 | 3.4 | 1960 | 7 | US-11-069-834-50 | Sequence 50, Appl |
| 17 | 163 | 3.4 | 5024 | 6 | US-10-793-626-2964 | Sequence 2964, Ap |
| 18 | 160.5 | 3.3 | 619 | 6 | US-10-485-517-374 | Sequence 374, App |
| 19 | 159 | 3.3 | 264 | 7 | US-11-188-743-24 | Sequence 24, Appl |
| 20 | 159 | 3.3 | 264 | 7 | US-11-188-743-25 | Sequence 25, Appl |
| 21 | 153.5 | 3.2 | 1960 | 7 | US-11-069-834-48 | Sequence 48, Appl |
| 22 | 152 | 3.2 | 663 | 7 | US-11-196-475-78 | Sequence 78, Appl |
| 23 | 149 | 3.1 | 663 | 7 | US-11-196-475-70 | Sequence 70, Appl |
| 24 | 148.5 | 3.1 | 693 | 7 | US-11-196-475-72 | Sequence 72, Appl |
| 25 | 147.5 | 3.1 | 1976 | 7 | US-11-069-834-52 | Sequence 52, Appl |
| 26 | 146.5 | 3.0 | 2096 | 6 | US-10-995-561-606 | Sequence 606, App |
| 27 | 146.5 | 3.0 | 2351 | 6 | US-10-995-561-608 | Sequence 608, App |
| 28 | 146 | 3.0 | 1279 | 6 | US-10-793-626-3188 | Sequence 3188, Ap |
| 29 | 145 | 3.0 | 416 | 6 | US-10-793-626-2 | Sequence 2, Appli |
| 30 | 143 | 3.0 | 1976 | 7 | US-11-069-834-54 | Sequence 54, Appl |
| 31 | 141 | 2.9 | 895 | 6 | US-10-485-517-129 | Sequence 129, App |
| 32 | 139.5 | 2.9 | 708 | 7 | US-11-196-475-76 | Sequence 76, Appl |
| 33 | 138.5 | 2.9 | 1147 | 6 | US-10-615-668-5 | Sequence 5, Appli |
| 34 | 138.5 | 2.9 | 1178 | 7 | US-11-044-899-29 | Sequence 29, Appl |
| 35 | 136.5 | 2.8 | 693 | 7 | US-11-196-475-68 | Sequence 68, Appl |
| 36 | 136.5 | 2.8 | 835 | 6 | US-10-501-039-4 | Sequence 4, Appli |
| 37 | 136.5 | 2.8 | 1095 | 6 | US-10-793-626-3154 | Sequence 3154, Ap |
| 38 | 135.5 | 2.8 | 700 | 7 | US-11-196-475-66 | Sequence 66, Appl |
| 39 | 135 | 2.8 | 1155 | 6 | US-10-793-626-1780 | Sequence 1780, Ap |
| 40 | 134.5 | 2.8 | 792 | 7 | US-11-108-172-1127 | Sequence 1127, Ap |
| 41 | 134.5 | 2.8 | 989 | 6 | US-10-821-234-975 | Sequence 975, App |
| 42 | 133.5 | 2.8 | 700 | 7 | US-11-196-475-74 | Sequence 74, Appl |
| 43 | 132.5 | 2.8 | 215 | 6 | US-10-821-234-1443 | Sequence 1443, Ap |
| 44 | 132.5 | 2.8 | 1107 | 6 | US-10-485-517-145 | Sequence 145, App |
| 45 | 130.5 | 2.7 | 1158 | 7 | US-11-075-646-6 | Sequence 6, Appli |

ALIGNMENTS

RESULT 1
US-11-188-743-17
; Sequence 17, Application US/11188743
; Publication No. US20050272140A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/11/188,743
; CURRENT FILING DATE: 2005-07-26
; PRIOR APPLICATION NUMBER: US/10/641,068
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US/09/788,657
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-188-743-17

Query Match 100.0%; Score 4812; DB 7; Length 932;
Best Local Similarity 100.0%; Pred. No. 6.5e-284;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
```

```
Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

```
RESULT 2
US-11-188-743-18
; Sequence 18, Application US/11188743
; Publication No. US20050272140A1
; GENERAL INFORMATION:
;  APPLICANT: Nicolaides, Nicholas
;  APPLICANT:  Sass, Philip
;  APPLICANT:  Kinzler, Kenneth
;  APPLICANT:  Grasso, Luigi
;  APPLICANT:  Vogelstein, Bert
;  TITLE OF INVENTION: Methods for generating hypermutable
;  TITLE OF INVENTION:  yeast
;  FILE REFERENCE: 01107.00097
;  CURRENT APPLICATION NUMBER: US/11/188,743
;  CURRENT FILING DATE:  2005-07-26
;  PRIOR APPLICATION NUMBER: US/10/641,068
;  PRIOR FILING DATE: 2003-08-15
;  PRIOR APPLICATION NUMBER: US/09/788,657
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: 60/184,336
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 18
;   LENGTH: 932
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-188-743-18

  Query Match            100.0%;  Score 4812;  DB 7;  Length 932;
  Best Local Similarity   100.0%;  Pred. No. 6.5e-284;
  Matches  932;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:20:47 ; Search time 189 Seconds
(without alignments)
2166.672 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 932
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 932 | 100.0 | 932 | 4 | AAB85852 | Aab85852 Human PMS |
| 2 | 932 | 100.0 | 932 | 4 | AAG63953 | Aag63953 Amino aci |
| 3 | 932 | 100.0 | 932 | 4 | AAG63954 | Aag63954 Amino aci |
| 4 | 932 | 100.0 | 932 | 5 | AAU98776 | Aau98776 Human pos |
| 5 | 932 | 100.0 | 932 | 5 | AAE24357 | Aae24357 Human mis |
| 6 | 932 | 100.0 | 932 | 5 | AAO18553 | Aao18553 Human mis |
| 7 | 932 | 100.0 | 932 | 5 | AAE28277 | Aae28277 Human MLH |
| 8 | 932 | 100.0 | 932 | 5 | AAE24684 | Aae24684 Human PMS |
| 9 | 932 | 100.0 | 932 | 6 | ABU07972 | Abu07972 Human PMS |
| 10 | 932 | 100.0 | 932 | 6 | ABU07971 | Abu07971 Human PMS |
| 11 | 932 | 100.0 | 932 | 6 | ABU89659 | Abu89659 Human PMS |
| 12 | 932 | 100.0 | 932 | 6 | ABU89660 | Abu89660 Human PMS |
| 13 | 932 | 100.0 | 932 | 6 | ABO07415 | Abo07415 Human Mut |
| 14 | 932 | 100.0 | 932 | 6 | AAO27514 | Aao27514 Human mis |
| 15 | 932 | 100.0 | 932 | 6 | AAO27515 | Aao27515 Human mis |
| 16 | 932 | 100.0 | 932 | 7 | ADA06246 | Ada06246 Human mis |
| 17 | 932 | 100.0 | 932 | 7 | ADA06244 | Ada06244 Human mis |
| 18 | 932 | 100.0 | 932 | 7 | ADC89605 | Adc89605 Human PMS |
| 19 | 932 | 100.0 | 932 | 7 | ADC89607 | Adc89607 Human PMS |
| 20 | 932 | 100.0 | 932 | 7 | ADF17892 | Adf17892 Human PMS |
| 21 | 932 | 100.0 | 932 | 7 | ADG62881 | Adg62881 Human PMS |
| 22 | 932 | 100.0 | 932 | 7 | ADG62882 | Adg62882 Human PMS |
| 23 | 932 | 100.0 | 932 | 7 | ADH62629 | Adh62629 Human mis |
| 24 | 932 | 100.0 | 932 | 7 | ADH60983 | Adh60983 Human mis |
| 25 | 932 | 100.0 | 932 | 7 | ADJ68675 | Adj68675 Human hea |
| 26 | 932 | 100.0 | 932 | 8 | ADF78840 | Adf78840 Human mis |
| 27 | 932 | 100.0 | 932 | 8 | ADG46767 | Adg46767 Human MMR |
| 28 | 932 | 100.0 | 932 | 8 | ADO40066 | Ado40066 Human PMS |
| 29 | 932 | 100.0 | 932 | 8 | ADP66682 | Adp66682 Human mis |
| 30 | 932 | 100.0 | 932 | 8 | ADP66680 | Adp66680 Human mis |
| 31 | 932 | 100.0 | 932 | 8 | ADR13883 | Adr13883 Human DNA |
| 32 | 932 | 100.0 | 932 | 8 | ADU77038 | Adu77038 Human mis |
| 33 | 932 | 100.0 | 932 | 9 | ADX08123 | Adx08123 Cyclin-de |
| 34 | 932 | 100.0 | 932 | 9 | ADX58473 | Adx58473 Amino aci |
| 35 | 932 | 100.0 | 932 | 9 | ADY53431 | Ady53431 Human PMS |
| 36 | 931 | 99.9 | 931 | 2 | AAR79009 | Aar79009 Human DNA |
| 37 | 900 | 96.6 | 900 | 4 | AAB85851 | Aab85851 Human PMS |
| 38 | 893 | 95.8 | 932 | 8 | ADT98686 | Adt98686 Human PMS |
| 39 | 893 | 95.8 | 932 | 8 | ADU77022 | Adu77022 Human mis |
| 40 | 696 | 74.7 | 847 | 8 | ABM83683 | Abm83683 Human dia |
| 41 | 232 | 24.9 | 232 | 5 | AAE28280 | Aae28280 Human MLH |
| 42 | 133 | 14.3 | 133 | 4 | AAB85855 | Aab85855 Human PMS |
| 43 | 133 | 14.3 | 133 | 4 | AAG63957 | Aag63957 Amino aci |
| 44 | 133 | 14.3 | 133 | 5 | AAO18559 | Aao18559 Human tru |
| 45 | 133 | 14.3 | 133 | 5 | AAE24687 | Aae24687 Human PMS |

ALIGNMENTS

RESULT 1
AAB85852
ID AAB85852 standard; protein; 932 AA.
XX
AC AAB85852;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PMS1 protein.
XX
KW Hypermutable bacteria; mismatch repair gene; MMR gene; MutH; MutS; MutL;
KW MutY; PMS2; MLH1; MLH3; PMSR; biocatalysis; bioremediation; biochemical;
KW drug discovery; detoxification; toxin; biotransformation; PMS1.
XX
OS Homo sapiens.
XX
PN WO200159092-A2.
XX
PD 16-AUG-2001.
XX
PF 12-FEB-2001; 2001WO-US004339.
XX
PR 11-FEB-2000; 2000US-0181929P.
XX
PA (UYJO ) UNIV JOHNS HOPKINS.
XX
PI Nicolaides NC, Sass PM, Grasso L, Vogelstein B, Kinzler KW;
XX
DR WPI; 2001-514664/56.
DR N-PSDB; AAH76365.
XX
PT Making hypermutable bacteria for biocatalysis, bioremediation and drug
PT discovery, involves introducing polynucleotide comprising dominant
PT negative allele of mismatch repair gene under regulatory sequence
PT control.
XX
PS Example 1; Page 41; 68pp; English.
XX
CC The invention provides a method for generating a hypermutable bacteria.
CC The method involves introducing a polynucleotide having a dominant
CC negative allele of a mismatch repair (MMR) gene under the control of an
CC inducible transcription regulatory sequence, into a bacterium. The cell
CC becomes inducibly hypermutable. The method is useful to create desirable
CC output traits for commercial applications, using dominant negative
CC alleles of mismatch repair proteins. The mismatch repair gene is a MutH,
CC MutS, MutL or MutY homologue and can be selected from PMS2, MLH1, MLH3,
CC PMSR or PMSR homologue. The hypermutable bacteria is useful for the

```
CC   production, biocatalysis, bioremediation and drug discovery. It is also
CC   useful in manufacturing industry for the generation of new biochemicals
CC   useful for detoxifying noxious chemicals from by-products of
CC   manufacturing processes or those used as catalysts, for remediation of
CC   toxins present in the environment including polychlorobenzenes, heavy
CC   metals and other environmental hazards for which there is a need to
CC   remove them from the environment. The hypermutable bacteria is also
CC   useful for screening novel mutations in a gene or a set of genes that
CC   produce variant siblings that exhibit a new output trait not found in
CC   wild type cells. The bacteria are also useful for producing increased
CC   quantity or quality of protein or non-protein therapeutic molecule e.g.
CC   Penicillin G, Erythromycin and Clavulanic acid, by biotransformation.
CC   Dominant negative alleles of the MMR gene are useful for producing higher
CC   quantities of recombinant polypeptides. The present sequence represents a
CC   human PMS1 protein
XX
SQ   Sequence 932 AA;

  Query Match             100.0%;  Score 932;  DB 4;  Length 932;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches  932;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932


RESULT 2
AAG63953
ID   AAG63953 standard; protein; 932 AA.
XX
AC   AAG63953;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Amino acid sequence of human mismatch repair protein PMS2.
XX
KW   PMS2; mismatch repair gene; MMR gene; hypermutable yeast.
XX
OS   Homo sapiens.
XX
PN   WO200162945-A1.
XX
PD   30-AUG-2001.
XX
PF   21-FEB-2001; 2001WO-US005447.
XX
PR   23-FEB-2000; 2000US-0184336P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
PA   (NICO/) NICOLAIDES N C.
PA   (SASS/) SASS P M.
PA   (GRAS/) GRASSO L.
PA   (VOGE/) VOGELSTEIN B.
PA   (KINZ/) KINZLER K W.
XX
PI   Nicolaides NC,  Sass PM,  Grasso L,  Vogelstein B,  Kinzler KW;
XX
DR   WPI; 2001-522820/57.
DR   N-PSDB; AAH75041.
XX
PT   Making hypermutable yeast that exhibit novel selected output traits for
PT   commercial applications, comprises introducing polynucleotide containing
PT   dominant negative allele of mismatch repair gene.
XX
PS   Disclosure; Page 38; 60pp; English.
XX
CC   The present sequence represents human PSM2. PMS2 is a mismatch repair
CC   (MMR) gene. The specification describes a method for making a
CC   hypermutable yeast, comprising introducing a polynucleotide containing a
CC   dominant negative allele of a mismatch repair (MMR) gene, into a yeast,
CC   whereby the cell becomes hypermutable. The method is useful to create
CC   desirable output traits for commercial applications, using dominant
CC   negative alleles of mismatch repair proteins. The hypermutable yeast is
CC   useful for production, biocatalysis, bioremediation and drug discovery.
CC   It is also useful in genetic screens for the direct selection of variant
CC   subclones that exhibit new output traits. The hypermutable yeast is also
CC   useful in the manufacturing industry for the generation of new
CC   biochemicals, for detoxifying noxious chemicals from by-products of
CC   manufacturing processes or those used as catalysts, for remediation of
CC   toxins present in the environment including polychlorobenzenes, heavy
CC   metals and other environmental hazards for which there is a need to
CC   remove them from the environment. The yeast is also useful for producing
CC   increased quantity or quality of protein or non-protein therapeutic
CC   molecule e.g., Penicillin G, Erythromycin and Clavulanic acid, by
CC   biotransformation
```

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:24:48 ; Search time 48 Seconds
(without alignments)
1605.286 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 932
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size : 0

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 932 | 100.0 | 932 | 2 | US-08-294-312B-4 | Sequence 4, Appli |
| 2 | 932 | 100.0 | 932 | 2 | US-08-468-024B-4 | Sequence 4, Appli |
| 3 | 932 | 100.0 | 932 | 2 | US-09-708-200-11 | Sequence 11, Appl |
| 4 | 932 | 100.0 | 932 | 2 | US-08-465-679-4 | Sequence 4, Appli |
| 5 | 932 | 100.0 | 932 | 2 | US-09-788-657-17 | Sequence 17, Appl |
| 6 | 932 | 100.0 | 932 | 2 | US-09-788-657-18 | Sequence 18, Appl |
| 7 | 932 | 100.0 | 932 | 2 | US-09-712-691-9 | Sequence 9, Appli |
| 8 | 932 | 100.0 | 932 | 2 | US-10-641-068-17 | Sequence 17, Appl |
| 9 | 932 | 100.0 | 932 | 2 | US-10-641-068-18 | Sequence 18, Appl |
| 10 | 912 | 97.9 | 921 | 2 | US-09-949-016-7787 | Sequence 7787, Ap |
| 11 | 893 | 95.8 | 932 | 2 | US-09-707-468C-9 | Sequence 9, Appli |
| 12 | 133 | 14.3 | 133 | 2 | US-09-788-657-21 | Sequence 21, Appl |
| 13 | 133 | 14.3 | 133 | 2 | US-09-712-691-15 | Sequence 15, Appl |
| 14 | 133 | 14.3 | 133 | 2 | US-10-641-068-21 | Sequence 21, Appl |
| 15 | 16 | 1.7 | 580 | 2 | US-09-198-452A-870 | Sequence 870, App |
| 16 | 16 | 1.7 | 594 | 2 | US-09-438-185A-814 | Sequence 814, App |
| 17 | 12 | 1.3 | 147 | 2 | US-09-749-601A-14 | Sequence 14, Appl |
| 18 | 12 | 1.3 | 426 | 2 | US-08-676-444-40 | Sequence 40, Appl |
| 19 | 12 | 1.3 | 779 | 2 | US-09-749-601A-12 | Sequence 12, Appl |
| 20 | 9 | 1.0 | 64 | 1 | US-08-209-521-19 | Sequence 19, Appl |
| 21 | 9 | 1.0 | 64 | 1 | US-08-209-521-20 | Sequence 20, Appl |
| 22 | 9 | 1.0 | 64 | 2 | US-08-961-810-129 | Sequence 129, App |
| 23 | 9 | 1.0 | 64 | 2 | US-08-961-810-130 | Sequence 130, App |
| 24 | 9 | 1.0 | 64 | 2 | US-08-352-902D-129 | Sequence 129, App |
| 25 | 9 | 1.0 | 64 | 2 | US-08-352-902D-130 | Sequence 130, App |
| 26 | 9 | 1.0 | 64 | 2 | US-09-265-503B-129 | Sequence 129, App |
| 27 | 9 | 1.0 | 64 | 2 | US-09-265-503B-130 | Sequence 130, App |
| 28 | 9 | 1.0 | 133 | 2 | US-09-708-200-17 | Sequence 17, Appl |
| 29 | 9 | 1.0 | 133 | 2 | US-09-707-468C-15 | Sequence 15, Appl |
| 30 | 9 | 1.0 | 133 | 2 | US-09-749-601A-13 | Sequence 13, Appl |
| 31 | 9 | 1.0 | 158 | 2 | US-09-489-039A-8515 | Sequence 8515, Ap |
| 32 | 9 | 1.0 | 244 | 2 | US-09-248-796A-19147 | Sequence 19147, A |
| 33 | 9 | 1.0 | 264 | 2 | US-09-788-657-24 | Sequence 24, Appl |
| 34 | 9 | 1.0 | 264 | 2 | US-09-788-657-25 | Sequence 25, Appl |
| 35 | 9 | 1.0 | 264 | 2 | US-10-641-068-24 | Sequence 24, Appl |
| 36 | 9 | 1.0 | 264 | 2 | US-10-641-068-25 | Sequence 25, Appl |
| 37 | 9 | 1.0 | 361 | 1 | US-08-209-521-5 | Sequence 5, Appli |
| 38 | 9 | 1.0 | 361 | 2 | US-08-961-810-1 | Sequence 1, Appli |
| 39 | 9 | 1.0 | 361 | 2 | US-08-352-902D-1 | Sequence 1, Appli |
| 40 | 9 | 1.0 | 361 | 2 | US-09-265-503B-1 | Sequence 1, Appli |
| 41 | 9 | 1.0 | 389 | 2 | US-09-788-657-23 | Sequence 23, Appl |
| 42 | 9 | 1.0 | 389 | 2 | US-10-641-068-23 | Sequence 23, Appl |
| 43 | 9 | 1.0 | 511 | 2 | US-08-676-444-42 | Sequence 42, Appl |
| 44 | 9 | 1.0 | 615 | 2 | US-08-676-444-44 | Sequence 44, Appl |
| 45 | 9 | 1.0 | 674 | 2 | US-09-543-681A-5715 | Sequence 5715, Ap |

ALIGNMENTS

RESULT 1
US-08-294-312B-4
; Sequence 4, Application US/08294312B
; Patent No. 6380369
; GENERAL INFORMATION:
; APPLICANT: Adams et al.
; TITLE OF INVENTION: Human DNA Mismatch Repair Proteins
; FILE REFERENCE: PF106P2
; CURRENT APPLICATION NUMBER: US/08/294,312B
; CURRENT FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/210,143
; PRIOR FILING DATE: 1994-03-16
; PRIOR APPLICATION NUMBER: 08/187,757
; PRIOR FILING DATE: 1994-01-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 932
; TYPE: PRT
; ORGANISM: homo sapiens
US-08-294-312B-4

Query Match 100.0%; Score 932; DB 2; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
```

```
Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
US-08-468-024B-4
; Sequence 4, Application US/08468024B
; Patent No. 6416984
; GENERAL INFORMATION:
; APPLICANT: Haseltine et al.
; TITLE OF INVENTION: Human DNA Mismatch Repair Proteins
; FILE REFERENCE: PF106P3
; CURRENT APPLICATION NUMBER: US/08/468,024B
; CURRENT FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/294,312
; PRIOR FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/210,143
; PRIOR FILING DATE: 1994-03-16
; PRIOR APPLICATION NUMBER: 08/187,757
; PRIOR FILING DATE: 1994-01-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 932
; TYPE: PRT
; ORGANISM: homo sapiens
US-08-468-024B-4

Query Match 100.0%; Score 932; DB 2; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 3
US-09-708-200-11
; Sequence 11, Application US/09708200
; Patent No. 6576468
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas C
; APPLICANT: Grasso, Luigi
; APPLICANT: Sass, Philip M
; TITLE OF INVENTION: METHODS FOR ISOLATING NOVEL ANTIMICROBIAL AGENTS FROM
; TITLE OF INVENTION: HYPERMUTABLE CELLS
; FILE REFERENCE: MOR-0005
; CURRENT APPLICATION NUMBER: US/09/708,200

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:32:28 ; Search time 168 Seconds
(without alignments)
2317.958 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 932
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1867569 seqs, 417829326 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA_Main:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 932 | 100.0 | 932 | 3 | US-09-788-657-17 | Sequence 17, Appl |
| 2 | 932 | 100.0 | 932 | 3 | US-09-788-657-18 | Sequence 18, Appl |
| 3 | 932 | 100.0 | 932 | 3 | US-09-912-697-6 | Sequence 6, Appli |
| 4 | 932 | 100.0 | 932 | 3 | US-09-912-697-8 | Sequence 8, Appli |
| 5 | 932 | 100.0 | 932 | 3 | US-09-760-285-18 | Sequence 18, Appl |
| 6 | 932 | 100.0 | 932 | 4 | US-10-079-429-4 | Sequence 4, Appli |
| 7 | 932 | 100.0 | 932 | 4 | US-10-270-839-27 | Sequence 27, Appl |
| 8 | 932 | 100.0 | 932 | 4 | US-10-270-839-29 | Sequence 29, Appl |
| 9 | 932 | 100.0 | 932 | 4 | US-10-243-130-9 | Sequence 9, Appli |
| 10 | 932 | 100.0 | 932 | 4 | US-10-371-857-3 | Sequence 3, Appli |
| 11 | 932 | 100.0 | 932 | 4 | US-10-371-857-4 | Sequence 4, Appli |
| 12 | 932 | 100.0 | 932 | 4 | US-10-371-634-5 | Sequence 5, Appli |
| 13 | 932 | 100.0 | 932 | 4 | US-10-371-634-7 | Sequence 7, Appli |
| 14 | 932 | 100.0 | 932 | 4 | US-10-348-074-3 | Sequence 3, Appli |
| 15 | 932 | 100.0 | 932 | 4 | US-10-348-074-5 | Sequence 5, Appli |
| 16 | 932 | 100.0 | 932 | 4 | US-10-369-845-11 | Sequence 11, Appl |
| 17 | 932 | 100.0 | 932 | 4 | US-10-641-068-17 | Sequence 17, Appl |
| 18 | 932 | 100.0 | 932 | 4 | US-10-641-068-18 | Sequence 18, Appl |
| 19 | 932 | 100.0 | 932 | 4 | US-10-408-765A-481 | Sequence 481, App |
| 20 | 932 | 100.0 | 932 | 4 | US-10-813-502-9 | Sequence 9, Appli |
| 21 | 932 | 100.0 | 932 | 4 | US-10-714-228-2 | Sequence 2, Appli |
| 22 | 932 | 100.0 | 932 | 4 | US-10-714-228-4 | Sequence 4, Appli |
| 23 | 932 | 100.0 | 932 | 5 | US-10-933-034-2 | Sequence 2, Appli |
| 24 | 932 | 100.0 | 932 | 5 | US-10-901-650-9 | Sequence 9, Appli |
| 25 | 932 | 100.0 | 932 | 6 | US-11-056-776-13 | Sequence 13, Appl |
| 26 | 893 | 95.8 | 932 | 5 | US-10-850-370-9 | Sequence 9, Appli |
| 27 | 133 | 14.3 | 133 | 3 | US-09-788-657-21 | Sequence 21, Appl |
| 28 | 133 | 14.3 | 133 | 3 | US-09-912-697-14 | Sequence 14, Appl |
| 29 | 133 | 14.3 | 133 | 3 | US-09-760-285-24 | Sequence 24, Appl |
| 30 | 133 | 14.3 | 133 | 4 | US-10-270-839-35 | Sequence 35, Appl |
| 31 | 133 | 14.3 | 133 | 4 | US-10-371-857-12 | Sequence 12, Appl |
| 32 | 133 | 14.3 | 133 | 4 | US-10-371-634-13 | Sequence 13, Appl |
| 33 | 133 | 14.3 | 133 | 4 | US-10-348-074-11 | Sequence 11, Appl |
| 34 | 133 | 14.3 | 133 | 4 | US-10-641-068-21 | Sequence 21, Appl |
| 35 | 133 | 14.3 | 133 | 4 | US-10-714-228-6 | Sequence 6, Appli |
| 36 | 16 | 1.7 | 580 | 4 | US-10-289-762-870 | Sequence 870, App |
| 37 | 16 | 1.7 | 580 | 4 | US-10-282-122A-54985 | Sequence 54985, A |
| 38 | 13 | 1.4 | 724 | 3 | US-09-954-950-2 | Sequence 2, Appli |
| 39 | 13 | 1.4 | 724 | 4 | US-10-270-839-59 | Sequence 59, Appl |
| 40 | 13 | 1.4 | 724 | 4 | US-10-714-228-50 | Sequence 50, Appl |
| 41 | 13 | 1.4 | 1141 | 4 | US-10-437-963-127061 | Sequence 127061, |
| 42 | 12 | 1.3 | 133 | 4 | US-10-714-228-38 | Sequence 38, Appl |
| 43 | 12 | 1.3 | 133 | 5 | US-10-933-034-38 | Sequence 38, Appl |
| 44 | 12 | 1.3 | 147 | 3 | US-09-749-601A-14 | Sequence 14, Appl |
| 45 | 12 | 1.3 | 779 | 3 | US-09-749-601A-12 | Sequence 12, Appl |

ALIGNMENTS

RESULT 1
US-09-788-657-17
; Sequence 17, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-657-17

Query Match 100.0%; Score 932; DB 3; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
```

```
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 2
US-09-788-657-18
; Sequence 18, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-788-657-18

Query Match 100.0%; Score 932; DB 3; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

RESULT 3
US-09-912-697-6
; Sequence 6, Application US/09912697
; Publication No. US20030068808A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas C
; APPLICANT: Sass, Philip M
; APPLICANT: Grasso, Luigi M

GenCore version 5.1.6


OM protein - protein search, using sw model

Run on: December 21, 2005, 20:33:19 ; Search time 13 Seconds
(without alignments)
511.317 Million cell updates/sec

Title: US-10-079-429A-4
Perfect score: 932
Sequence: 1 MKQLPAATVRLLSSSQIITS.....KECVHGRPFFHHLTYLPETT 932

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 53982 seqs, 7132107 residues

Word size : 0

Total number of hits satisfying chosen parameters: 53982

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_AA_New:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 932 | 100.0 | 932 | 7 | US-11-188-743-17 | Sequence 17, Appl |
| 2 | 932 | 100.0 | 932 | 7 | US-11-188-743-18 | Sequence 18, Appl |
| 3 | 133 | 14.3 | 133 | 7 | US-11-188-743-21 | Sequence 21, Appl |
| 4 | 12 | 1.3 | 147 | 7 | US-11-128-420-14 | Sequence 14, Appl |
| 5 | 12 | 1.3 | 779 | 7 | US-11-128-420-12 | Sequence 12, Appl |
| 6 | 9 | 1.0 | 133 | 7 | US-11-128-420-13 | Sequence 13, Appl |
| 7 | 9 | 1.0 | 264 | 7 | US-11-188-743-24 | Sequence 24, Appl |
| 8 | 9 | 1.0 | 264 | 7 | US-11-188-743-25 | Sequence 25, Appl |
| 9 | 9 | 1.0 | 389 | 7 | US-11-188-743-23 | Sequence 23, Appl |
| 10 | 9 | 1.0 | 756 | 7 | US-11-188-743-20 | Sequence 20, Appl |
| 11 | 9 | 1.0 | 769 | 7 | US-11-188-743-15 | Sequence 15, Appl |
| 12 | 9 | 1.0 | 862 | 7 | US-11-128-420-11 | Sequence 11, Appl |
| 13 | 8 | 0.9 | 859 | 7 | US-11-188-743-16 | Sequence 16, Appl |
| 14 | 8 | 0.9 | 1151 | 7 | US-11-128-420-10 | Sequence 10, Appl |
| 15 | 7 | 0.8 | 11 | 6 | US-10-839-799-114 | Sequence 114, App |
| 16 | 7 | 0.8 | 11 | 6 | US-10-839-799-129 | Sequence 129, App |
| 17 | 7 | 0.8 | 11 | 7 | US-11-108-135-7 | Sequence 7, Appli |
| 18 | 7 | 0.8 | 20 | 7 | US-11-040-159-4 | Sequence 4, Appli |
| 19 | 7 | 0.8 | 20 | 7 | US-11-054-669-107 | Sequence 107, App |
| 20 | 7 | 0.8 | 28 | 7 | US-11-174-089-152 | Sequence 152, App |
| 21 | 7 | 0.8 | 113 | 7 | US-11-144-248-20 | Sequence 20, Appl |
| 22 | 7 | 0.8 | 114 | 7 | US-11-065-943-47 | Sequence 47, Appl |
| 23 | 7 | 0.8 | 114 | 7 | US-11-055-163-18 | Sequence 18, Appl |
| 24 | 7 | 0.8 | 116 | 7 | US-11-174-186-20 | Sequence 20, Appl |
| 25 | 7 | 0.8 | 116 | 7 | US-11-174-186-21 | Sequence 21, Appl |
| 26 | 7 | 0.8 | 116 | 7 | US-11-174-186-22 | Sequence 22, Appl |
| 27 | 7 | 0.8 | 116 | 7 | US-11-174-186-23 | Sequence 23, Appl |
| 28 | 7 | 0.8 | 116 | 7 | US-11-174-186-24 | Sequence 24, Appl |
| 29 | 7 | 0.8 | 116 | 7 | US-11-174-186-25 | Sequence 25, Appl |
| 30 | 7 | 0.8 | 116 | 7 | US-11-055-163-17 | Sequence 17, Appl |
| 31 | 7 | 0.8 | 117 | 6 | US-10-839-799-132 | Sequence 132, App |
| 32 | 7 | 0.8 | 117 | 7 | US-11-054-669-120 | Sequence 120, App |
| 33 | 7 | 0.8 | 118 | 6 | US-10-932-334-75 | Sequence 75, Appl |
| 34 | 7 | 0.8 | 118 | 7 | US-11-012-353-71 | Sequence 71, Appl |
| 35 | 7 | 0.8 | 118 | 7 | US-11-009-939-12 | Sequence 12, Appl |
| 36 | 7 | 0.8 | 120 | 6 | US-10-932-334-71 | Sequence 71, Appl |
| 37 | 7 | 0.8 | 120 | 7 | US-11-096-074-2 | Sequence 2, Appli |
| 38 | 7 | 0.8 | 120 | 7 | US-11-009-939-2 | Sequence 2, Appli |
| 39 | 7 | 0.8 | 120 | 7 | US-11-107-028-17 | Sequence 17, Appl |
| 40 | 7 | 0.8 | 121 | 7 | US-11-108-135-24 | Sequence 24, Appl |
| 41 | 7 | 0.8 | 121 | 7 | US-11-127-677-54 | Sequence 54, Appl |
| 42 | 7 | 0.8 | 121 | 7 | US-11-107-028-16 | Sequence 16, Appl |
| 43 | 7 | 0.8 | 121 | 7 | US-11-107-028-18 | Sequence 18, Appl |
| 44 | 7 | 0.8 | 121 | 7 | US-11-107-028-19 | Sequence 19, Appl |
| 45 | 7 | 0.8 | 121 | 7 | US-11-107-028-20 | Sequence 20, Appl |

ALIGNMENTS

RESULT 1
US-11-188-743-17
; Sequence 17, Application US/11188743
; Publication No. US20050272140A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/11/188,743
; CURRENT FILING DATE: 2005-07-26
; PRIOR APPLICATION NUMBER: US/10/641,068
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US/09/788,657
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 932
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-188-743-17

Query Match 100.0%; Score 932; DB 7; Length 932;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 932; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
```

```
Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

```
RESULT 2
US-11-188-743-18
; Sequence 18, Application US/11188743
; Publication No. US20050272140A1
; GENERAL INFORMATION:
;  APPLICANT: Nicolaides, Nicholas
;  APPLICANT:  Sass, Philip
;  APPLICANT:  Kinzler, Kenneth
;  APPLICANT:  Grasso, Luigi
;  APPLICANT:  Vogelstein, Bert
;  TITLE OF INVENTION: Methods for generating hypermutable
;  TITLE OF INVENTION:  yeast
;  FILE REFERENCE: 01107.00097
;  CURRENT APPLICATION NUMBER: US/11/188,743
;  CURRENT FILING DATE:  2005-07-26
;  PRIOR APPLICATION NUMBER: US/10/641,068
;  PRIOR FILING DATE: 2003-08-15
;  PRIOR APPLICATION NUMBER: US/09/788,657
;  PRIOR FILING DATE: 2001-02-21
;  PRIOR APPLICATION NUMBER: 60/184,336
;  PRIOR FILING DATE: 2000-02-23
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 18
;   LENGTH: 932
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-11-188-743-18

  Query Match           100.0%;  Score 932;  DB 7;  Length 932;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches  932;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRDNGEG 60

Qy         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 IKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTRTAADNFSTQ 120

Qy        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 YVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKCKDEIKKIQDLLMSFG 180

Qy        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVMNNMESFQYHSEESQIYLSGFL 240

Qy        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 PKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIRHHYNLKCLKESTRLYPVFFLKID 300

Qy        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 VPTADVDVNLTPDKSQVLLQNKESVLIALENLMTTCYGPLPSTNSYENNKTDVSAADIVL 360

Qy        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 SKTAETDVLFNKVESSGKNYSNVDTSVIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGH 420

Qy        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CSSEISNIDKNTKNAFQDISMSNVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGE 480

Qy        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 NEEEAGLENSSEISADEWSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMN 540

Qy        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 LNEDSCNKKSNVIDNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLED 600

Qy        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATLQIEELWKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAW 660

Qy        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 NLAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKVDLEE 720

Qy        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 KDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTE 780

Qy        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 SLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMAN 840

Qy        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDI 900

Qy        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
              ||||||||||||||||||||||||||||||||
Db        901 IYRMKHQFGNEIKECVHGRPFFHHLTYLPETT 932
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 05:00:57 ; Search time 10058 Seconds
(without alignments)
17310.747 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 6 | BD181098 | BD181098 Human DNA |
| 2 | 3063 | 100.0 | 3063 | 6 | CS020326 | CS020326 Sequence |
| 3 | 3063 | 100.0 | 3063 | 6 | CS056082 | CS056082 Sequence |
| 4 | 3063 | 100.0 | 3063 | 6 | AR217665 | AR217665 Sequence |
| 5 | 3063 | 100.0 | 3063 | 6 | AR342820 | AR342820 Sequence |
| 6 | 3063 | 100.0 | 3063 | 6 | AR380890 | AR380890 Sequence |
| 7 | 3063 | 100.0 | 3063 | 6 | AR382023 | AR382023 Sequence |
| 8 | 3063 | 100.0 | 3063 | 6 | AR437149 | AR437149 Sequence |
| 9 | 3063 | 100.0 | 3063 | 6 | AR540769 | AR540769 Sequence |
| 10 | 3063 | 100.0 | 3063 | 6 | AR592873 | AR592873 Sequence |
| 11 | 3063 | 100.0 | 3063 | 6 | AX214170 | AX214170 Sequence |
| 12 | 3063 | 100.0 | 3063 | 6 | AX234586 | AX234586 Sequence |
| 13 | 3063 | 100.0 | 3063 | 6 | AX775137 | AX775137 Sequence |
| 14 | 3063 | 100.0 | 3063 | 8 | HSU13695 | U13695 Human homol |
| 15 | 3051.4 | 99.6 | 3239 | 6 | CS130851 | CS130851 Sequence |
| 16 | 2830 | 92.4 | 2830 | 8 | BC096332 | BC096332 Homo sapi |
| 17 | 2826.8 | 92.3 | 2830 | 8 | BC096330 | BC096330 Homo sapi |
| 18 | 2555 | 83.4 | 2682 | 8 | AB102875 | AB102875 Homo sapi |
| 19 | 2482 | 81.0 | 2856 | 6 | CQ720889 | CQ720889 Sequence |
| 20 | 2239 | 73.1 | 2817 | 8 | AY540751 | AY540751 Homo sapi |
| 21 | 2016.6 | 65.8 | 2252 | 8 | AY540750 | AY540750 Homo sapi |
| 22 | 1912.4 | 62.4 | 2343 | 8 | BC096331 | BC096331 Homo sapi |
| 23 | 1887 | 61.6 | 3045 | 9 | BC028939 | BC028939 Mus muscu |
| 24 | 1856 | 60.6 | 2313 | 8 | AB102870 | AB102870 Homo sapi |
| 25 | 1848.2 | 60.3 | 3009 | 9 | BC061722 | BC061722 Rattus no |
| 26 | 1845 | 60.2 | 2004 | 8 | AB102869 | AB102869 Homo sapi |
| 27 | 1222.6 | 39.9 | 3147 | 5 | AJ719999 | AJ719999 Gallus ga |
| 28 | 1160.2 | 37.9 | 1668 | 8 | AB102872 | AB102872 Homo sapi |
| 29 | 896.6 | 29.3 | 3276 | 5 | BC044098 | BC044098 Xenopus l |
| 30 | 894.2 | 29.2 | 104140 | 8 | AC008122 | AC008122 Homo sapi |
| 31 | 894.2 | 29.2 | 106234 | 8 | AY267352 | AY267352 Homo sapi |
| 32 | 891 | 29.1 | 155122 | 14 | AC141847 | AC141847 Pan trogl |
| c 33 | 891 | 29.1 | 212131 | 14 | AC142554 | AC142554 Pan trogl |
| 34 | 857.4 | 28.0 | 188715 | 14 | AC141850 | AC141850 Papio anu |
| 35 | 849 | 27.7 | 2894 | 5 | BC089718 | BC089718 Xenopus t |
| 36 | 816 | 26.6 | 184340 | 14 | AC155201 | AC155201 Callithri |
| 37 | 701.4 | 22.9 | 705 | 8 | AB102876 | AB102876 Homo sapi |
| 38 | 584.8 | 19.1 | 747 | 8 | AB102874 | AB102874 Homo sapi |
| 39 | 583.2 | 19.0 | 591 | 8 | AB102873 | AB102873 Homo sapi |
| 40 | 582 | 19.0 | 588 | 8 | AB102877 | AB102877 Homo sapi |
| 41 | 563.6 | 18.4 | 2058 | 5 | CR762073 | CR762073 Xenopus t |
| 42 | 555.8 | 18.1 | 994 | 5 | CR353667 | CR353667 Gallus ga |
| 43 | 496 | 16.2 | 1906 | 8 | BC008410 | BC008410 Homo sapi |
| 44 | 492.6 | 16.1 | 1229 | 5 | BX929478 | BX929478 Gallus ga |
| 45 | 489.4 | 16.0 | 965 | 8 | BC084548 | BC084548 Homo sapi |

ALIGNMENTS

RESULT 1
BD181098
LOCUS BD181098 3063 bp DNA linear PAT 15-MAY-2003
DEFINITION Human DNA mismatch repair proteins.
ACCESSION BD181098
VERSION BD181098.1 GI:30792016
KEYWORDS JP 2002325588-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3063)
AUTHORS Haseltine,W.A., Ruben,S.M., Wei,Y.F., Adams,M.D., Fleischmann,R.D., Fraser,C.M., Fuldner,R.A., Kirkness,E.F. and Rosen,C.A.
TITLE Human DNA mismatch repair proteins
JOURNAL Patent: JP 2002325588-A 2 12-NOV-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002325588-A/2
PD 12-NOV-2002
PF 25-JAN-2002 JP 2002016830
PR 27-JAN-1994 US 08/187757,16-MAR-1994 US 08/210143 PR 23-AUG-1994 US 08/294312
PI WILLIAM A HASELTINE,STEVEN M RUBEN,YING FEI WEI,MARK D ADAMS,
PI ROBERT D FLEISCHMANN,CLAIRE M FRASER,REBECCA A FULDNER,EWEN F
PI KIRKNESS,
PI CRAIG A ROSEN
PC C12N15/09,C07K14/47,C12P21/02,C12Q1/68//(C12P21/02,C12R1:19),
PC C12N15/00
CC Human DNA mismatch repair proteins
FH Key Location/Qualifiers
FT CDS (81). .(2879).
FEATURES Location/Qualifiers
source 1. .3063
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 3063; DB 6; Length 3063;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300

Qy        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360

Qy        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420

Qy        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480

Qy        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540

Qy        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600

Qy        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660

Qy        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720

Qy        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780

Qy        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840

Qy        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900

Qy        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960

Qy        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020

Qy       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080

Qy       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140

Qy       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200

Qy       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260

Qy       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320

Qy       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380

Qy       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440

Qy       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500

Qy       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560

Qy       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620

Qy       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680

Qy       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740

Qy       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800

Qy       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860

Qy       1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920

Qy       1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980

Qy       1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040

Qy       2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100

Qy       2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 00:25:02 ; Search time 1173 Seconds
(without alignments)
17403.200 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 2 | AAQ97526 | Aaq97526 Human DNA |
| 2 | 3063 | 100.0 | 3063 | 4 | AAH76365 | Aah76365 Human PMS |
| 3 | 3063 | 100.0 | 3063 | 5 | AAH75042 | Aah75042 Nucleotid |
| 4 | 3063 | 100.0 | 3063 | 6 | ABK86089 | Abk86089 Human cDN |
| 5 | 3063 | 100.0 | 3063 | 6 | AAD39198 | Aad39198 Human mis |
| 6 | 3063 | 100.0 | 3063 | 6 | AAL48698 | Aal48698 Human mis |
| 7 | 3063 | 100.0 | 3063 | 6 | AAD45354 | Aad45354 Human MLH |
| 8 | 3063 | 100.0 | 3063 | 6 | AAD39770 | Aad39770 Human PMS |
| 9 | 3063 | 100.0 | 3063 | 8 | ABX12940 | Abx12940 DNA encod |
| 10 | 3063 | 100.0 | 3063 | 8 | ACA89704 | Aca89704 cDNA enco |
| 11 | 3063 | 100.0 | 3063 | 9 | ACD13762 | Acd13762 Human Mut |
| 12 | 3063 | 100.0 | 3063 | 9 | AAL57764 | Aal57764 Human mis |
| 13 | 3063 | 100.0 | 3063 | 9 | ADA06245 | Ada06245 DNA encod |
| 14 | 3063 | 100.0 | 3063 | 10 | ADC89608 | Adc89608 Human PMS |
| 15 | 3063 | 100.0 | 3063 | 10 | ADE85236 | Ade85236 Farnesyl |
| 16 | 3063 | 100.0 | 3063 | 10 | ADF17893 | Adf17893 Human PMS |
| 17 | 3063 | 100.0 | 3063 | 10 | ADG62892 | Adg62892 Human PMS |
| 18 | 3063 | 100.0 | 3063 | 10 | ADH62630 | Adh62630 Human mis |
| 19 | 3063 | 100.0 | 3063 | 10 | ADH60982 | Adh60982 Human cDN |
| 20 | 3063 | 100.0 | 3063 | 11 | ADI32109 | Adi32109 Human cDN |
| 21 | 3063 | 100.0 | 3063 | 12 | ADF78841 | Adf78841 Human mis |
| 22 | 3063 | 100.0 | 3063 | 12 | ADG46768 | Adg46768 Human MMR |
| 23 | 3063 | 100.0 | 3063 | 12 | ADO40067 | Ado40067 Human PMS |
| 24 | 3063 | 100.0 | 3063 | 12 | ADP66679 | Adp66679 Human mis |
| 25 | 3063 | 100.0 | 3063 | 13 | ADR13884 | Adr13884 Human DNA |
| 26 | 3063 | 100.0 | 3063 | 13 | ADT98687 | Adt98687 Human PMS |
| 27 | 3063 | 100.0 | 3063 | 13 | ADU77023 | Adu77023 Human mis |
| 28 | 3063 | 100.0 | 3063 | 13 | ADS84176 | Ads84176 Human lym |
| 29 | 3063 | 100.0 | 3063 | 14 | ADX58474 | Adx58474 Nucleotid |
| 30 | 3063 | 100.0 | 3063 | 14 | ADY53430 | Ady53430 Human PMS |
| 31 | 3061.4 | 99.9 | 3063 | 6 | AAD45433 | Aad45433 Human MLH |
| 32 | 3051.4 | 99.6 | 3239 | 14 | AEB22842 | Aeb22842 Human col |
| 33 | 2991 | 97.6 | 3121 | 14 | ADX08122 | Adx08122 Cyclin-de |
| 34 | 2486.6 | 81.2 | 2793 | 13 | ACN42335 | Acn42335 Human dia |
| 35 | 421 | 13.7 | 493 | 13 | ADQ79340 | Adq79340 Novel can |
| 36 | 297.2 | 9.7 | 327 | 8 | ABX12507 | Abx12507 DNA repai |
| 37 | 297.2 | 9.7 | 327 | 9 | ADA06170 | Ada06170 Human EST |
| 38 | 276.8 | 9.0 | 559 | 5 | ABV48030 | Abv48030 Human pro |
| 39 | 180 | 5.9 | 189 | 2 | AAT19265 | Aat19265 Human gen |
| 40 | 178.4 | 5.8 | 184 | 3 | AAC22447 | Aac22447 Human sec |
| 41 | 171 | 5.6 | 534 | 5 | ABV18243 | Abv18243 Human pro |
| 42 | 126 | 4.1 | 478 | 9 | ACH23591 | Ach23591 Human adu |
| 43 | 112.4 | 3.7 | 478 | 13 | ADR29677 | Adr29677 Mouse gen |
| 44 | 101.2 | 3.3 | 2772 | 8 | ACA89712 | Aca89712 cDNA enco |
| 45 | 101.2 | 3.3 | 2772 | 12 | ADP66717 | Adp66717 A. thalia |

ALIGNMENTS

RESULT 1
AAQ97526
ID AAQ97526 standard; cDNA; 3063 BP.
XX
AC AAQ97526;
XX
DT 27-MAR-1996 (first entry)
XX
DE Human DNA repair protein hMLH2 coding sequence.
XX
KW DNA repair protein; hMLH1; hMLH2; hMLH3; therapy; cancer; vectors;
KW DNA synthesis; diagnosis; disease; mutL4; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 81. .2873
FT /*tag= a
FT /product= "DNA repair protein."
FT /note= "The CDS may extend to position 2879 where a TGA
FT stop codon is located. There could possibly be an error
FT in the decoded protein in the specification since, if the
FT stop codon is the correct translation termination signal,
FT the C-terminal end of the protein should end with two
FT Threonine residues. It is possible that one of these has
FT been omitted."
XX
PN WO9520678-A1.
XX
PD 03-AUG-1995.
XX
PF 25-JAN-1995; 95WO-US001035.
XX
PR 27-JAN-1994; 94US-00187757.
PR 16-MAR-1994; 94US-00210143.
PR 23-AUG-1994; 94US-00294312.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haseltine WA, Ruben SM, Wei Y, Adams MD, Fleischmann RD;
PI Fraser CM, Fuldner RA, Kirkness EF, Rosen CA;
XX

DR WPI; 1995-275461/36.
DR P-PSDB; AAR79009.
XX
PT Polynucleotide(s) encoding human mutL4 homologues, hMLH1, hMLH2 and hMLH3
PT - used for therapeutic treatment of, e.g. hereditary cancer.
XX
PS Claim 1; Fig 2; 124pp; English.
XX
CC The polynucleotides described in AAQ97525-27 encode the human analogues
CC of the prokaryotic mutL4 DNA repair gene. The polypeptides they encode
CC (AAR79008-R79010) are used for therapeutic purposes e.g. in the treatment
CC of cancer, esp. hereditary cancer. They may also be used for in vitro
CC manipulation of DNA, synthesis of DNA and the manufacture of DNA vectors
CC and in methods of diagnosing a disease or a susceptibilty to a disease
CC related to a mutation in the hMLH1, -2 or -3 DNA repair genes
XX
SQ Sequence 3063 BP; 1100 A; 503 C; 580 G; 880 T; 0 U; 0 Other;

Query Match 100.0%; Score 3063; DB 2; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300

Qy        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360

Qy        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420

Qy        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480

Qy        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540

Qy        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600

Qy        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660

Qy        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720

Qy        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780

Qy        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840

Qy        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900

Qy        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960

Qy        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020

Qy       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080

Qy       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140

Qy       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200

Qy       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260

Qy       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320

Qy       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380

Qy       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440

Qy       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500

Qy       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560

Qy       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620

Qy       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680

Qy       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740

Qy       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800

Qy       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 05:14:07 ; Search time 7732 Seconds
(without alignments)
18534.510 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 2969.8 | 97.0 | 3130 | 4 | CR859593 | CR859593 Pongo pyg |
| | 2 | 2164.2 | 70.7 | 2365 | 4 | BC036376 | BC036376 Homo sapi |
| | 3 | 1970 | 64.3 | 2667 | 11 | DQ052952 | DQ052952 Homo sapi |
| | 4 | 1825.8 | 59.6 | 2667 | 11 | DQ052953 | DQ052953 Pan trogl |
| | 5 | 1804 | 58.9 | 1804 | 4 | CR610658 | CR610658 full-leng |
| | 6 | 891.2 | 29.1 | 1011 | 5 | BX353664 | BX353664 BX353664 |
| | 7 | 846 | 27.6 | 993 | 7 | CO580445 | CO580445 ILLUMIGEN |
| c | 8 | 835.4 | 27.3 | 893 | 5 | BX328949 | BX328949 BX328949 |
| | 9 | 821.8 | 26.8 | 903 | 5 | BX327629 | BX327629 BX327629 |
| | 10 | 819.2 | 26.7 | 1021 | 3 | BM479838 | BM479838 AGENCOURT |
| c | 11 | 803.6 | 26.2 | 898 | 5 | BX435290 | BX435290 BX435290 |
| | 12 | 780.6 | 25.5 | 1117 | 3 | BM553209 | BM553209 AGENCOURT |
| | 13 | 780 | 25.5 | 780 | 7 | CN336501 | CN336501 170006001 |
| | 14 | 776 | 25.3 | 776 | 5 | BX117693 | BX117693 BX117693 |
| | 15 | 721 | 23.5 | 733 | 7 | CN336498 | CN336498 170006001 |
| | 16 | 718 | 23.4 | 851 | 8 | DN102367 | DN102367 1097287 M |
| | 17 | 717.8 | 23.4 | 857 | 1 | AM019977 | AM019977 AM019977 |
| | 18 | 708.2 | 23.1 | 820 | 5 | BX419980 | BX419980 BX419980 |
| | 19 | 706.2 | 23.1 | 801 | 2 | BG193431 | BG193431 RST12563 |
| c | 20 | 704.6 | 23.0 | 986 | 5 | BX353663 | BX353663 BX353663 |
| | 21 | 702.8 | 22.9 | 842 | 8 | DN526212 | DN526212 1270959 M |
| | 22 | 700.2 | 22.9 | 956 | 3 | BM800196 | BM800196 AGENCOURT |
| c | 23 | 694 | 22.7 | 722 | 3 | BM677668 | BM677668 UI-E-EO1- |
| c | 24 | 677 | 22.1 | 704 | 5 | BQ771615 | BQ771615 UI-H-EZ1- |
| c | 25 | 676.6 | 22.1 | 902 | 8 | DN523518 | DN523518 1267023 M |
| | 26 | 650.2 | 21.2 | 694 | 1 | AL705101 | AL705101 DKFZp686H |
| | 27 | 640 | 20.9 | 641 | 6 | CB157373 | CB157373 K-EST0216 |
| c | 28 | 639.4 | 20.9 | 748 | 6 | CA415469 | CA415469 UI-H-EZ0- |
| c | 29 | 637.6 | 20.8 | 706 | 6 | CB852984 | CB852984 UI-CF-FN0 |
| c | 30 | 634.4 | 20.7 | 746 | 5 | BU623174 | BU623174 UI-H-FL1- |
| c | 31 | 634 | 20.7 | 638 | 1 | AL043809 | AL043809 DKFZp434P |
| | 32 | 633.2 | 20.7 | 835 | 2 | BG168340 | BG168340 602342315 |
| | 33 | 627 | 20.5 | 753 | 5 | BX925658 | BX925658 BX925658 |
| | 34 | 626.4 | 20.5 | 668 | 8 | DN998411 | DN998411 TC106042 |
| | 35 | 621 | 20.3 | 843 | 3 | BI464618 | BI464618 603203232 |
| | 36 | 620.8 | 20.3 | 782 | 5 | BQ429685 | BQ429685 AGENCOURT |
| | 37 | 617.2 | 20.2 | 949 | 5 | BX452128 | BX452128 BX452128 |
| | 38 | 616.4 | 20.1 | 852 | 2 | BG536475 | BG536475 602564576 |
| | 39 | 613.8 | 20.0 | 825 | 3 | BI545790 | BI545790 603187943 |
| | 40 | 603.8 | 19.7 | 735 | 5 | BX925660 | BX925660 BX925660 |
| | 41 | 602 | 19.7 | 603 | 3 | BM723144 | BM723144 UI-E-EO1- |
| | 42 | 601.8 | 19.6 | 869 | 7 | CO737632 | CO737632 S1LH03c18 |
| | 43 | 601.8 | 19.6 | 988 | 2 | BG163660 | BG163660 602338840 |
| | 44 | 598.2 | 19.5 | 905 | 2 | BF666456 | BF666456 602123905 |
| c | 45 | 596 | 19.5 | 623 | 2 | BE350913 | BE350913 ht63h02.x |

ALIGNMENTS

RESULT 1
CR859593
LOCUS CR859593 3130 bp mRNA linear HTC 12-NOV-2004
DEFINITION Pongo pygmaeus mRNA; cDNA DKFZp468M105 (from clone DKFZp468M105).
ACCESSION CR859593
VERSION CR859593.1 GI:55730062
KEYWORDS HTC.
SOURCE Pongo pygmaeus (orangutan)
ORGANISM Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Pongo.
REFERENCE 1 (bases 1 to 3130)
AUTHORS Ansorge,W., Krieger,S., Regiert,T., Rittmueller,C., Schwager,B., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
CONSRTM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp468M105) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp468M105
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES Location/Qualifiers
source 1. .3130
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp468M105"
/tissue_type="heart"
/clone_lib="468 (synonym: phrt1). Vector pSport1_Sfi; host DH10B; sites SfilA + SfilB"
/dev_stage="adult"
/note="PMS1 protein homolog 1 (Homo sapiens)"
gene 1. .3130
/gene="DKFZp468M105"
CDS 135. .2930

```
                     /gene="DKFZp468M105"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAH91756.1"
                     /db_xref="GI:55730063"
                     /translation="MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSIDVKL
                     ENYGFDKIEVRDNGEGIKAVDTPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICC
                     VAEVLITTRTAADNFSTQYVLDGSGHILSQKPSHLAQGTTVTALRLFRNLPVRKQFYS
                     TAKKCKDEIKKIQDLLMSYGILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAV
                     MSNMESFQYHSEESQIYLSGFLPKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLI
                     RHHYNLKCLKESTRLYPVFFLKIDVPTADVDVNLTPDKSQVLLQNKESVLIALENLMT
                     TCYGPLLSTNSYENNKTDVSAADIVLSKTAETDVLFNKMESSGKNYSNVDTSIIPFQN
                     DMHNDESGKNNDDCLNHQISIGDFGYGHCSSEISNVDKNTKNAFQDISMSNVSWENTQ
                     TEYSKTCFVGSIKHTQSENGNKDHIDESGENEEEAGLENSLEISADEWSRGNILKNSV
                     GENIEPVKILVPEKSLPCKVSNNNYPSPEQTNLNEDPCNKKSNVIDNKSGKVTAYDLL
                     SNRVVKKPMSASALFVQDHRPQFLIENPKTSLEDATLQIEELWKTLSEEEKLKYEEKA
                     TKDLERYNNQMKRAIEQESQISLKDGRKKIKPTSAWNLAQKHKLKTSLSNQPKLDELL
                     QSQIEKRSQNIKMVQIPFSMKNLKINFKKQDKVDLEEKNEPCLIHNLRFPDAWLMTSK
                     TEVMLLNPYRVEEALLFKRLLENHKLPAEPLEKPIMLTESLFNGSHYLDVLYKMTADD
                     QRYSGSTYLSDPRLTANGFKIKLIPGVSITENYLEIEGMANCLPFYGVADLKEILNAI
                     LNKNAKEVYECRPRKVISYLEGEAVRLSRQLPMYLSKEDIQDIIYRMKHQFGNEIKGC
                     VHGRPFFHHLTYLPETT"
ORIGIN

  Query Match             97.0%;  Score 2969.8;  DB 4;  Length 3130;
  Best Local Similarity   98.5%;  Pred. No. 0;
  Matches 3008;  Conservative    0;  Mismatches   42;  Indels    3;  Gaps    1;

Qy          9 TGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAGCTGCTCTG 68
              |||||||||||| |||||||||||||||||||| |||||||||||| || ||||||||||
Db         63 TGGCTGCTTGCCGCTAGTGGATGGTAATTGCTTGCCTCGCGCTAGAAGGAAGCTGCTCTG 122

Qy         69 TTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAAGTTCTCAG 128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 TTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAAGTTCTCAG 182

Qy        129 ATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGGATGCTGGT 188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 ATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGGATGCTGGT 242

Qy        189 GCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGGTGCGAGAT 248
              |||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        243 GCCACAAGCATAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGGTGCGAGAT 302

Qy        249 AACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACTACACCTCA 308
              |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db        303 AACGGGGAGGGTATCAAGGCTGTTGATACACCTGTAATGGCAATGAAGTACTACACCTCA 362

Qy        309 AAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTGGAGAAGCC 368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        363 AAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTGGAGAAGCC 422

Qy        369 TTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTGCTGATAAT 428
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db        423 TTGGGGTCAATTTGTTGTGTAGCTGAGGTTTTAATTACAACAAGAACGGCTGCTGATAAT 482

Qy        429 TTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAACCTTCACAT 488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        483 TTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAACCTTCACAT 542

Qy        489 CTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTGTAAGAAAG 548
              |||| ||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db        543 CTTGCTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAGGAATCTACCTGTAAGAAAG 602

Qy        549 CAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAGATCTCCTC 608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        603 CAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAGATCTCCTC 662

Qy        609 ATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACAAGGCAGTT 668
              ||||||| ||||||||||||||||||||||||||||| || ||||||||||||||||||
Db        663 ATGAGCTATGGTATCCTTAAACCTGACTTAAGGATTGTTTTCGTACATAACAAGGCAGTT 722

Qy        669 ATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTCTGGGGACT 728
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        723 ATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTCTGGGGACT 782

Qy        729 GCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGATTTATCTC 788
              |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 GCTGTTATGAGCAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGATTTATCTC 842

Qy        789 AGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAACACCAGAA 848
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        843 AGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAACACCAGAA 902

Qy        849 AGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAAAGTTAATC 908
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        903 AGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAAAGTTAATC 962

Qy        909 CGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTGTTTTCTTT 968
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        963 CGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTGTTTTCTTT 1022

Qy        969 CTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATAAAAGCCAA 1028
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1023 CTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATAAAAGTCAA 1082

Qy       1029 GTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGACGACTTGT 1088
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1083 GTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGACGACTTGT 1142

Qy       1089 TATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTTCCGCAGCT 1148
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db       1143 TATGGACCATTACTTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTTCCGCAGCT 1202

Qy       1149 GACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGGAATCATCT 1208
              |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db       1203 GACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAATGGAATCATCT 1262

Qy       1209 GGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATATGCATAAT 1268
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1263 GGAAAGAATTATTCAAATGTTGATACTTCAATCATTCCATTCCAAAATGATATGCATAAT 1322

Qy       1269 GATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTGGTGACTTT 1328
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db       1323 GATGAATCTGGAAAAAACAATGATGATTGTTTAAATCACCAGATAAGTATTGGTGACTTT 1382

Qy       1329 GGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGAATGCATTT 1388
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1383 GGTTATGGTCATTGTAGTAGTGAAATTTCTAACGTTGATAAAAACACTAAGAATGCATTT 1442

Qy       1389 CAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATAGTAAAACT 1448
              || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1443 CAAGACATTTCAATGAGTAATGTATCATGGGAGAACACTCAGACGGAATATAGTAAAACT 1502

Qy       1449 TGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACCATATAGAT 1508
              |||||| || |||||||||||||||||||||||||||||||||||||||||||||||||
Db       1503 TGTTTTGTAGGTTCCATTAAGCACACCCAGTCAGAAAATGGCAATAAAGACCATATAGAT 1562

Qy       1509 GAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTTCTGCAGAT 1568
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db       1563 GAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTTGGAAATTTCTGCAGAT 1622

Qy       1569 GAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAACCTGTGAAA 1628
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1623 GAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAACCTGTGAAA 1682

Qy       1629 ATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATCCAATCCCT 1688
              |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db       1683 ATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATCCAAGCCCT 1742

Qy       1689 GAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAGATAATAAA 1748
              ||||||| |||||||||||||||| ||||||||||||||||||||||||||||||||||
Db       1743 GAACAAACGAATCTTAATGAAGATCCATGTAACAAAAAATCAAATGTAATAGATAATAAA 1802

Qy       1749 TCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAACCCATGTCA 1808
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 06:20:02 ; Search time 371 Seconds
(without alignments)
14675.664 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 3 | US-08-294-312B-3 | Sequence 3, Appli |
| 2 | 3063 | 100.0 | 3063 | 3 | US-08-468-024B-3 | Sequence 3, Appli |
| 3 | 3063 | 100.0 | 3063 | 3 | US-09-708-200-12 | Sequence 12, Appl |
| 4 | 3063 | 100.0 | 3063 | 3 | US-09-023-655-1435 | Sequence 1435, Ap |
| 5 | 3063 | 100.0 | 3063 | 3 | US-08-465-679-3 | Sequence 3, Appli |
| 6 | 3063 | 100.0 | 3063 | 3 | US-09-788-657-8 | Sequence 8, Appli |
| 7 | 3063 | 100.0 | 3063 | 3 | US-09-712-691-10 | Sequence 10, Appl |
| 8 | 3063 | 100.0 | 3063 | 3 | US-09-707-468C-10 | Sequence 10, Appl |
| 9 | 3063 | 100.0 | 3063 | 3 | US-10-641-068-8 | Sequence 8, Appli |
| 10 | 2808.4 | 91.7 | 2817 | 3 | US-09-949-016-1916 | Sequence 1916, Ap |
| 11 | 894.2 | 29.2 | 96845 | 3 | US-09-949-016-13658 | Sequence 13658, A |
| 12 | 297.2 | 9.7 | 327 | 3 | US-08-187-757D-3 | Sequence 3, Appli |
| 13 | 297.2 | 9.7 | 327 | 3 | US-08-210-143C-3 | Sequence 3, Appli |
| 14 | 178.4 | 5.8 | 184 | 3 | US-09-513-999C-26522 | Sequence 26522, A |
| 15 | 134 | 4.4 | 601 | 3 | US-09-949-016-65618 | Sequence 65618, A |
| 16 | 99.8 | 3.3 | 2589 | 3 | US-09-749-601A-3 | Sequence 3, Appli |
| 17 | 99.8 | 3.3 | 2687 | 2 | US-08-209-521-22 | Sequence 22, Appl |
| 18 | 99.8 | 3.3 | 2687 | 3 | US-08-961-810-132 | Sequence 132, App |
| 19 | 99.8 | 3.3 | 2687 | 3 | US-08-352-902D-132 | Sequence 132, App |
| 20 | 99.8 | 3.3 | 2687 | 3 | US-09-265-503B-132 | Sequence 132, App |
| 21 | 99.8 | 3.3 | 2771 | 3 | US-09-059-461-1 | Sequence 1, Appli |
| 22 | 99.8 | 3.3 | 2771 | 3 | US-08-294-312B-5 | Sequence 5, Appli |
| 23 | 99.8 | 3.3 | 2771 | 3 | US-08-468-024B-5 | Sequence 5, Appli |
| 24 | 99.8 | 3.3 | 2771 | 3 | US-09-708-200-10 | Sequence 10, Appl |
| 25 | 99.8 | 3.3 | 2771 | 3 | US-08-465-679-5 | Sequence 5, Appli |
| 26 | 99.8 | 3.3 | 2771 | 3 | US-09-788-657-7 | Sequence 7, Appli |
| 27 | 99.8 | 3.3 | 2771 | 3 | US-09-712-691-8 | Sequence 8, Appli |
| 28 | 99.8 | 3.3 | 2771 | 3 | US-09-707-468C-8 | Sequence 8, Appli |
| 29 | 99.8 | 3.3 | 2771 | 3 | US-10-641-068-7 | Sequence 7, Appli |
| 30 | 96.2 | 3.1 | 821 | 3 | US-08-352-902D-146 | Sequence 146, App |
| 31 | 94.2 | 3.1 | 1408 | 3 | US-09-788-657-12 | Sequence 12, Appl |
| 32 | 94.2 | 3.1 | 1408 | 3 | US-10-641-068-12 | Sequence 12, Appl |
| 33 | 91.4 | 3.0 | 3056 | 3 | US-09-708-200-8 | Sequence 8, Appli |
| 34 | 91.4 | 3.0 | 3056 | 3 | US-09-788-657-6 | Sequence 6, Appli |
| 35 | 91.4 | 3.0 | 3056 | 3 | US-09-712-691-6 | Sequence 6, Appli |
| 36 | 91.4 | 3.0 | 3056 | 3 | US-09-707-468C-6 | Sequence 6, Appli |
| 37 | 91.4 | 3.0 | 3056 | 3 | US-10-641-068-6 | Sequence 6, Appli |
| 38 | 91.4 | 3.0 | 3065 | 2 | US-08-209-521-28 | Sequence 28, Appl |
| 39 | 91.4 | 3.0 | 3065 | 3 | US-09-265-503B-137 | Sequence 137, App |
| 40 | 85.4 | 2.8 | 654 | 3 | US-09-248-796A-5045 | Sequence 5045, Ap |
| 41 | 84.8 | 2.8 | 402 | 3 | US-09-749-601A-5 | Sequence 5, Appli |
| 42 | 84 | 2.7 | 426 | 3 | US-09-708-200-18 | Sequence 18, Appl |
| 43 | 84 | 2.7 | 426 | 3 | US-09-788-657-11 | Sequence 11, Appl |
| 44 | 84 | 2.7 | 426 | 3 | US-09-712-691-16 | Sequence 16, Appl |
| 45 | 84 | 2.7 | 426 | 3 | US-09-707-468C-16 | Sequence 16, Appl |

ALIGNMENTS

RESULT 1
US-08-294-312B-3
; Sequence 3, Application US/08294312B
; Patent No. 6380369
; GENERAL INFORMATION:
; APPLICANT: Adams et al.
; TITLE OF INVENTION: Human DNA Mismatch Repair Proteins
; FILE REFERENCE: PF106P2
; CURRENT APPLICATION NUMBER: US/08/294,312B
; CURRENT FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/210,143
; PRIOR FILING DATE: 1994-03-16
; PRIOR APPLICATION NUMBER: 08/187,757
; PRIOR FILING DATE: 1994-01-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3063
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2879)
US-08-294-312B-3

Query Match 100.0%; Score 3063; DB 3; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Qy      301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db      301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Qy      361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db      361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Qy      421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db      421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Qy      481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db      481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Qy      541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db      541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Qy      601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db      601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Qy      661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db      661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Qy      721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db      721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Qy      781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db      781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Qy      841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db      841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Qy      901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db      901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Qy      961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db      961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Qy     1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db     1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Qy     1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db     1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Qy     1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db     1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Qy     1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db     1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Qy     1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db     1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Qy     1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db     1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
```

```
Qy     1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db     1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Qy     1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db     1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Qy     1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db     1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Qy     1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db     1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Qy     1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db     1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Qy     1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db     1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Qy     1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db     1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Qy     1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db     1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Qy     1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Db     1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Qy     1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Db     1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Qy     1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Db     1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Qy     2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Db     2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Qy     2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Db     2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Qy     2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Db     2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Qy     2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Db     2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Qy     2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Db     2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Qy     2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Db     2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Qy     2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460
Db     2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 09:28:31 ; Search time 1567 Seconds
(without alignments)
16164.075 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 3 | US-09-788-657-8 | Sequence 8, Appli |
| 2 | 3063 | 100.0 | 3063 | 3 | US-09-912-697-7 | Sequence 7, Appli |
| 3 | 3063 | 100.0 | 3063 | 3 | US-09-760-285-19 | Sequence 19, Appl |
| 4 | 3063 | 100.0 | 3063 | 5 | US-10-079-429-3 | Sequence 3, Appli |
| 5 | 3063 | 100.0 | 3063 | 6 | US-10-270-839-28 | Sequence 28, Appl |
| 6 | 3063 | 100.0 | 3063 | 6 | US-10-243-130-10 | Sequence 10, Appl |
| 7 | 3063 | 100.0 | 3063 | 6 | US-10-371-857-14 | Sequence 14, Appl |
| 8 | 3063 | 100.0 | 3063 | 6 | US-10-371-634-8 | Sequence 8, Appli |
| 9 | 3063 | 100.0 | 3063 | 6 | US-10-348-074-6 | Sequence 6, Appli |
| 10 | 3063 | 100.0 | 3063 | 6 | US-10-369-845-12 | Sequence 12, Appl |
| 11 | 3063 | 100.0 | 3063 | 7 | US-10-641-643-1435 | Sequence 1435, Ap |
| 12 | 3063 | 100.0 | 3063 | 7 | US-10-641-068-8 | Sequence 8, Appli |
| 13 | 3063 | 100.0 | 3063 | 7 | US-10-283-975A-453 | Sequence 453, App |
| 14 | 3063 | 100.0 | 3063 | 7 | US-10-813-502-10 | Sequence 10, Appl |
| 15 | 3063 | 100.0 | 3063 | 8 | US-10-714-228-1 | Sequence 1, Appli |
| 16 | 3063 | 100.0 | 3063 | 8 | US-10-850-370-10 | Sequence 10, Appl |
| 17 | 3063 | 100.0 | 3063 | 9 | US-10-933-034-1 | Sequence 1, Appli |
| 18 | 3063 | 100.0 | 3063 | 9 | US-10-901-650-10 | Sequence 10, Appl |
| 19 | 3063 | 100.0 | 3063 | 10 | US-11-056-776-14 | Sequence 14, Appl |
| 20 | 2799 | 91.4 | 2799 | 6 | US-10-109-791A-9 | Sequence 9, Appli |
| 21 | 610.6 | 19.9 | 611 | 4 | US-09-925-065A-562210 | Sequence 562210, |
| 22 | 366.2 | 12.0 | 419 | 7 | US-10-242-535A-25367 | Sequence 25367, A |
| 23 | 366.2 | 12.0 | 419 | 7 | US-10-085-783A-25367 | Sequence 25367, A |
| 24 | 276.8 | 9.0 | 559 | 8 | US-10-357-930-48049 | Sequence 48049, A |
| 25 | 171 | 5.6 | 534 | 8 | US-10-357-930-18234 | Sequence 18234, A |
| 26 | 126 | 4.1 | 478 | 3 | US-09-918-995-10803 | Sequence 10803, A |
| 27 | 112.4 | 3.7 | 478 | 7 | US-10-702-075-254 | Sequence 254, App |
| 28 | 101.2 | 3.3 | 2772 | 6 | US-10-270-839-44 | Sequence 44, Appl |
| 29 | 101.2 | 3.3 | 2772 | 8 | US-10-714-228-39 | Sequence 39, Appl |
| 30 | 101.2 | 3.3 | 3099 | 8 | US-10-714-228-35 | Sequence 35, Appl |
| 31 | 101.2 | 3.3 | 3099 | 9 | US-10-933-034-35 | Sequence 35, Appl |
| 32 | 99.8 | 3.3 | 2589 | 3 | US-09-749-601A-3 | Sequence 3, Appli |
| 33 | 99.8 | 3.3 | 2589 | 6 | US-10-109-791A-10 | Sequence 10, Appl |
| 34 | 99.8 | 3.3 | 2656 | 9 | US-10-450-763-21391 | Sequence 21391, A |
| 35 | 99.8 | 3.3 | 2687 | 6 | US-10-349-607-132 | Sequence 132, App |
| 36 | 99.8 | 3.3 | 2771 | 3 | US-09-788-657-7 | Sequence 7, Appli |
| 37 | 99.8 | 3.3 | 2771 | 3 | US-09-912-697-5 | Sequence 5, Appli |
| 38 | 99.8 | 3.3 | 2771 | 3 | US-09-760-285-17 | Sequence 17, Appl |
| 39 | 99.8 | 3.3 | 2771 | 5 | US-10-079-429-5 | Sequence 5, Appli |
| 40 | 99.8 | 3.3 | 2771 | 6 | US-10-270-839-26 | Sequence 26, Appl |
| 41 | 99.8 | 3.3 | 2771 | 6 | US-10-243-130-8 | Sequence 8, Appli |
| 42 | 99.8 | 3.3 | 2771 | 6 | US-10-371-857-13 | Sequence 13, Appl |
| 43 | 99.8 | 3.3 | 2771 | 6 | US-10-371-634-6 | Sequence 6, Appli |
| 44 | 99.8 | 3.3 | 2771 | 6 | US-10-348-074-4 | Sequence 4, Appli |
| 45 | 99.8 | 3.3 | 2771 | 6 | US-10-369-845-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-09-788-657-8
; Sequence 8, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 3063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-788-657-8

Query Match 100.0%; Score 3063; DB 3; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

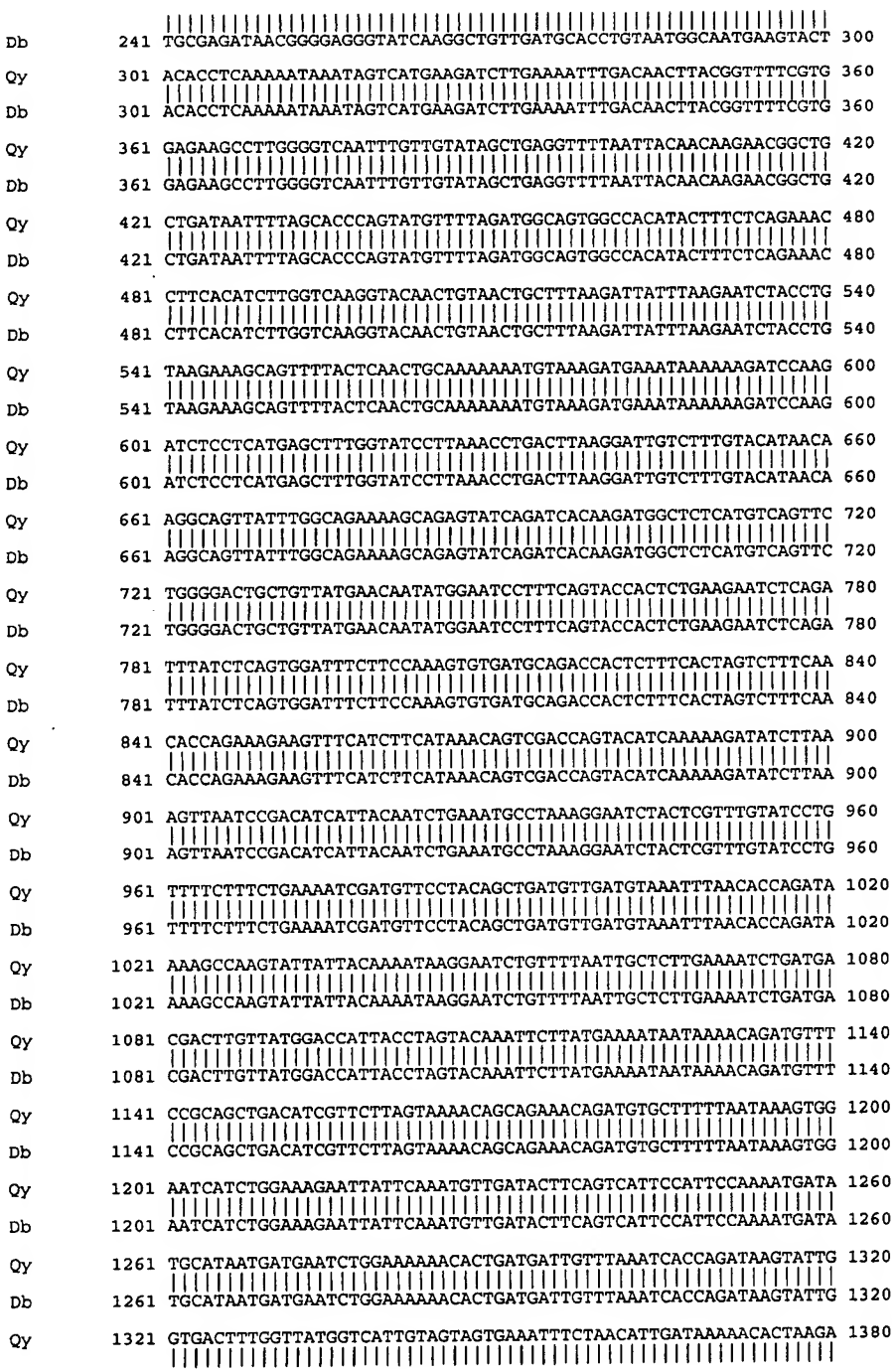
```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
```


Db 241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Qy 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Qy 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Qy 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Qy 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Qy 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Qy 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Qy 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Qy 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Qy 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Qy 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Qy 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Qy 961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db 961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Qy 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Qy 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Qy 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Qy 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Qy 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Qy 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Qy 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Qy 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Qy 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Qy 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Qy 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Qy 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Qy 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Qy 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Qy 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Db 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Qy 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Db 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Qy 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Db 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Qy 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Db 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Qy 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Db 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Qy 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Db 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Qy 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Db 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Qy 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Db 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Qy 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Db 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Qy 2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460
Db 2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 11:37:28 ; Search time 244 Seconds
(without alignments)
6514.585 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4168288 seqs, 259477437 residues

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 3063 | 100.0 | 3063 | 7 | US-11-188-743-8 | Sequence 8, Appli |
| | 2 | 99.8 | 3.3 | 2589 | 7 | US-11-128-420-3 | Sequence 3, Appli |
| | 3 | 99.8 | 3.3 | 2771 | 7 | US-11-188-743-7 | Sequence 7, Appli |
| | 4 | 94.2 | 3.1 | 1408 | 7 | US-11-188-743-12 | Sequence 12, Appl |
| | 5 | 91.4 | 3.0 | 3056 | 7 | US-11-188-743-6 | Sequence 6, Appli |
| | 6 | 84.8 | 2.8 | 402 | 7 | US-11-128-420-5 | Sequence 5, Appli |
| | 7 | 84 | 2.7 | 426 | 7 | US-11-188-743-11 | Sequence 11, Appl |
| | 8 | 80.4 | 2.6 | 1785 | 7 | US-11-188-743-13 | Sequence 13, Appl |
| | 9 | 76.8 | 2.5 | 2484 | 7 | US-11-188-743-10 | Sequence 10, Appl |
| | 10 | 75.2 | 2.5 | 3218 | 7 | US-11-188-743-5 | Sequence 5, Appli |
| | 11 | 72.4 | 2.4 | 2340 | 7 | US-11-128-420-4 | Sequence 4, Appli |
| | 12 | 55 | 1.8 | 3000 | 7 | US-11-194-246-150 | Sequence 150, App |
| | 13 | 53.4 | 1.7 | 441 | 7 | US-11-128-420-6 | Sequence 6, Appli |
| | 14 | 53 | 1.7 | 1974 | 6 | US-10-467-657-4781 | Sequence 4781, Ap |
| | 15 | 52 | 1.7 | 795 | 7 | US-11-188-743-14 | Sequence 14, Appl |
| | 16 | 49.4 | 1.6 | 1284 | 7 | US-11-074-176-343 | Sequence 343, App |
| | 17 | 49.4 | 1.6 | 1299 | 7 | US-11-074-176-161 | Sequence 161, App |
| c | 18 | 48.6 | 1.6 | 134499 | 7 | US-11-117-187-192 | Sequence 192, App |
| | 19 | 45.8 | 1.5 | 305312 | 6 | US-10-995-561-13236 | Sequence 13236, A |
| | 20 | 43 | 1.4 | 190276 | 6 | US-10-661-966-1 | Sequence 1, Appli |
| c | 21 | 42.6 | 1.4 | 2088 | 6 | US-10-793-626-3463 | Sequence 3463, Ap |
| c | 22 | 42.4 | 1.4 | 191797 | 7 | US-11-121-086-13 | Sequence 13, Appl |
| | 23 | 41.6 | 1.4 | 600 | 6 | US-10-750-185-20965 | Sequence 20965, A |
| | 24 | 41.6 | 1.4 | 1514 | 6 | US-10-750-185-52668 | Sequence 52668, A |
| c | 25 | 41.6 | 1.4 | 398287 | 6 | US-10-995-561-13396 | Sequence 13396, A |
| | 26 | 41.4 | 1.4 | 861 | 6 | US-10-793-626-2321 | Sequence 2321, Ap |
| | 27 | 41.4 | 1.4 | 3240 | 6 | US-10-793-626-3743 | Sequence 3743, Ap |
| c | 28 | 41 | 1.3 | 139054 | 7 | US-11-121-086-96 | Sequence 96, Appl |
| c | 29 | 40.8 | 1.3 | 863 | 6 | US-10-750-185-47174 | Sequence 47174, A |
| | 30 | 40.8 | 1.3 | 1625 | 6 | US-10-750-185-63003 | Sequence 63003, A |
| c | 31 | 40.8 | 1.3 | 1878 | 6 | US-10-750-185-57736 | Sequence 57736, A |
| | 32 | 40.6 | 1.3 | 894 | 6 | US-10-750-185-52936 | Sequence 52936, A |
| c | 33 | 40.6 | 1.3 | 2563 | 6 | US-10-485-517-399 | Sequence 399, App |
| c | 34 | 40.6 | 1.3 | 3146 | 6 | US-10-793-626-3375 | Sequence 3375, Ap |
| | 35 | 40.4 | 1.3 | 2389 | 6 | US-10-793-626-4438 | Sequence 4438, Ap |
| | 36 | 40.4 | 1.3 | 2619 | 6 | US-10-750-185-42549 | Sequence 42549, A |
| c | 37 | 40.4 | 1.3 | 87672 | 6 | US-10-995-561-13237 | Sequence 13237, A |
| | 38 | 40.2 | 1.3 | 648 | 6 | US-10-821-234-591 | Sequence 591, App |
| | 39 | 40.2 | 1.3 | 142303 | 7 | US-11-121-086-42 | Sequence 42, Appl |
| | 40 | 40 | 1.3 | 130472 | 6 | US-10-995-561-13312 | Sequence 13312, A |
| | 41 | 39.8 | 1.3 | 900 | 6 | US-10-793-626-1887 | Sequence 1887, Ap |
| | 42 | 39.6 | 1.3 | 5556 | 6 | US-10-955-054A-38 | Sequence 38, Appl |
| | 43 | 39.4 | 1.3 | 1356 | 6 | US-10-750-185-56485 | Sequence 56485, A |
| c | 44 | 39.4 | 1.3 | 4224 | 6 | US-10-750-185-56253 | Sequence 56253, A |
| c | 45 | 39.4 | 1.3 | 164527 | 7 | US-11-121-086-71 | Sequence 71, Appl |

ALIGNMENTS

RESULT 1
US-11-188-743-8
; Sequence 8, Application US/11188743
; Publication No. US20050272140A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/11/188,743
; CURRENT FILING DATE: 2005-07-26
; PRIOR APPLICATION NUMBER: US/10/641,068
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US/09/788,657
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 3063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-188-743-8

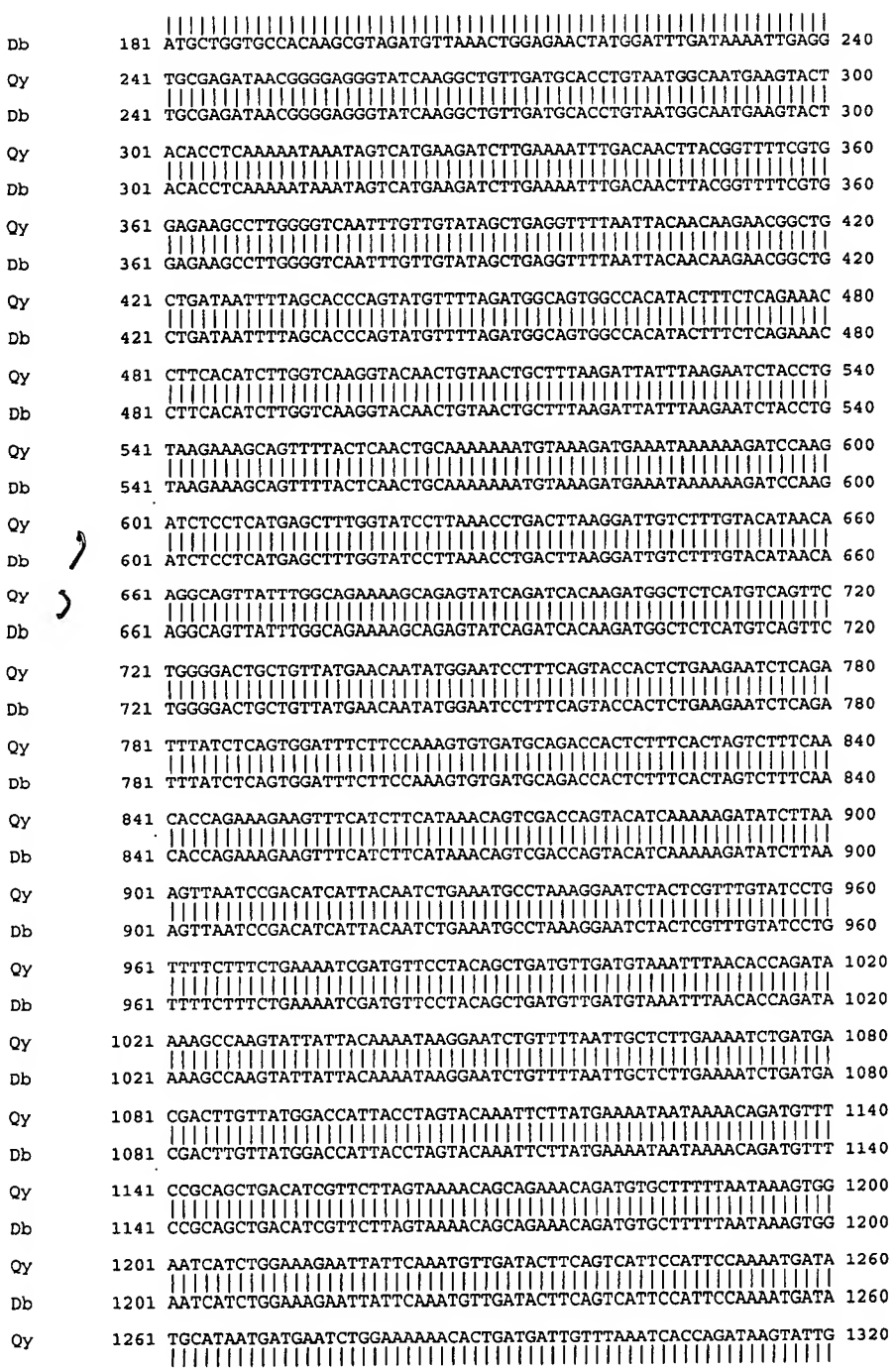
Query Match 100.0%; Score 3063; DB 7; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
```


Db 181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
Qy 241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Db 241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Qy 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Qy 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Qy 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Qy 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Qy 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Qy 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Qy 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Qy 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Qy 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Qy 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Qy 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Qy 961 TTTTCTTTCTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db 961 TTTTCTTTCTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Qy 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Qy 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Qy 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Qy 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Qy 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Qy 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Qy 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Qy 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Qy 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Qy 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Qy 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Qy 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Qy 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Qy 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Qy 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Db 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Qy 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Db 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Qy 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Db 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Qy 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Db 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Qy 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Db 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Qy 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Db 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Qy 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Db 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Qy 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Db 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Qy 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Db 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 14:40:21 ; Search time 10058 Seconds
(without alignments)
17310.747 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 5883141 seqs, 28421725653 residues

Word size : 0

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 3063 | 100.0 | 3063 | 6 | BD181098 | BD181098 Human DNA |
| | 2 | 3063 | 100.0 | 3063 | 6 | CS020326 | CS020326 Sequence |
| | 3 | 3063 | 100.0 | 3063 | 6 | CS056082 | CS056082 Sequence |
| | 4 | 3063 | 100.0 | 3063 | 6 | AR217665 | AR217665 Sequence |
| | 5 | 3063 | 100.0 | 3063 | 6 | AR342820 | AR342820 Sequence |
| | 6 | 3063 | 100.0 | 3063 | 6 | AR380890 | AR380890 Sequence |
| | 7 | 3063 | 100.0 | 3063 | 6 | AR382023 | AR382023 Sequence |
| | 8 | 3063 | 100.0 | 3063 | 6 | AR437149 | AR437149 Sequence |
| | 9 | 3063 | 100.0 | 3063 | 6 | AR540769 | AR540769 Sequence |
| | 10 | 3063 | 100.0 | 3063 | 6 | AR592873 | AR592873 Sequence |
| | 11 | 3063 | 100.0 | 3063 | 6 | AX214170 | AX214170 Sequence |
| | 12 | 3063 | 100.0 | 3063 | 6 | AX234586 | AX234586 Sequence |
| | 13 | 3063 | 100.0 | 3063 | 6 | AX775137 | AX775137 Sequence |
| | 14 | 3063 | 100.0 | 3063 | 8 | HSU13695 | U13695 Human homol |
| | 15 | 3004 | 98.1 | 3239 | 6 | CS130851 | CS130851 Sequence |
| | 16 | 2830 | 92.4 | 2830 | 8 | BC096332 | BC096332 Homo sapi |
| | 17 | 2732 | 89.2 | 2830 | 8 | BC096330 | BC096330 Homo sapi |
| | 18 | 2102 | 68.6 | 2682 | 8 | AB102875 | AB102875 Homo sapi |
| | 19 | 1979 | 64.6 | 2856 | 6 | CQ720889 | CQ720889 Sequence |
| | 20 | 1865 | 60.9 | 2343 | 8 | BC096331 | BC096331 Homo sapi |
| | 21 | 1856 | 60.6 | 2313 | 8 | AB102870 | AB102870 Homo sapi |
| | 22 | 1624 | 53.0 | 2252 | 8 | AY540750 | AY540750 Homo sapi |
| | 23 | 1392 | 45.4 | 2004 | 8 | AB102869 | AB102869 Homo sapi |
| | 24 | 1391 | 45.4 | 2817 | 8 | AY540751 | AY540751 Homo sapi |
| | 25 | 1159 | 37.8 | 1668 | 8 | AB102872 | AB102872 Homo sapi |
| | 26 | 892 | 29.1 | 104140 | 8 | AC008122 | AC008122 Homo sapi |
| | 27 | 892 | 29.1 | 106234 | 8 | AY267352 | AY267352 Homo sapi |
| | 28 | 790 | 25.8 | 155122 | 14 | AC141847 | AC141847 Pan trogl |
| c | 29 | 790 | 25.8 | 212131 | 14 | AC142554 | AC142554 Pan trogl |
| | 30 | 701 | 22.9 | 705 | 8 | AB102876 | AB102876 Homo sapi |
| | 31 | 583 | 19.0 | 747 | 8 | AB102874 | AB102874 Homo sapi |
| | 32 | 582 | 19.0 | 588 | 8 | AB102877 | AB102877 Homo sapi |
| | 33 | 582 | 19.0 | 591 | 8 | AB102873 | AB102873 Homo sapi |
| | 34 | 493 | 16.1 | 1906 | 8 | BC008410 | BC008410 Homo sapi |
| | 35 | 486 | 15.9 | 731 | 10 | G67580 | G67580 csnppms1-pc |
| | 36 | 442 | 14.4 | 965 | 8 | BC084548 | BC084548 Homo sapi |
| | 37 | 421 | 13.7 | 501 | 8 | BT006947 | BT006947 Homo sapi |
| | 38 | 421 | 13.7 | 501 | 11 | AY335564 | AY335564 Synthetic |
| | 39 | 421 | 13.7 | 501 | 11 | BT007647 | BT007647 Synthetic |
| | 40 | 233 | 7.6 | 419 | 6 | CQ680441 | CQ680441 Sequence |
| | 41 | 211 | 6.9 | 188715 | 14 | AC141850 | AC141850 Papio anu |
| | 42 | 187 | 6.1 | 327 | 6 | AR255689 | AR255689 Sequence |
| | 43 | 187 | 6.1 | 327 | 6 | AR399511 | AR399511 Sequence |
| | 44 | 167 | 5.5 | 389 | 10 | G67582 | G67582 csnppms1-pc |
| | 45 | 154 | 5.0 | 559 | 6 | CQ516182 | CQ516182 Sequence |

ALIGNMENTS

RESULT 1
BD181098
LOCUS BD181098 3063 bp DNA linear PAT 15-MAY-2003
DEFINITION Human DNA mismatch repair proteins.
ACCESSION BD181098
VERSION BD181098.1 GI:30792016
KEYWORDS JP 2002325588-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3063)
AUTHORS Haseltine,W.A., Ruben,S.M., Wei,Y.F., Adams,M.D., Fleischmann,R.D., Fraser,C.M., Fuldner,R.A., Kirkness,E.F. and Rosen,C.A.
TITLE Human DNA mismatch repair proteins
JOURNAL Patent: JP 2002325588-A 2 12-NOV-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002325588-A/2
PD 12-NOV-2002
PF 25-JAN-2002 JP 2002016830
PR 27-JAN-1994 US 08/187757,16-MAR-1994 US 08/210143 PR 23-AUG-1994 US 08/294312
PI WILLIAM A HASELTINE,STEVEN M RUBEN,YING FEI WEI,MARK D ADAMS,
PI ROBERT D FLEISCHMANN,CLAIRE M FRASER,REBECCA A FULDNER,EWEN F
PI KIRKNESS,
PI CRAIG A ROSEN
PC C12N15/09,C07K14/47,C12P21/02,C12Q1/68//(C12P21/02,C12R1:19),
PC C12N15/00
CC Human DNA mismatch repair proteins
FH Key Location/Qualifiers
FT CDS (81)..(2879).
FEATURES Location/Qualifiers
source 1..3063
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 3063; DB 6; Length 3063;

```
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 3063;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300

Qy        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360

Qy        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420

Qy        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480

Qy        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540

Qy        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600

Qy        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660

Qy        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720

Qy        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780

Qy        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840

Qy        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900

Qy        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960

Qy        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020

Qy       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080

Qy       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140

Qy       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200

Qy       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260

Qy       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320

Qy       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380

Qy       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440

Qy       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500

Qy       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560

Qy       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620

Qy       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680

Qy       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740

Qy       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800

Qy       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860

Qy       1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920

Qy       1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980

Qy       1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040

Qy       2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100

Qy       2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 14:34:00 ; Search time 1173 Seconds
(without alignments)
17403.200 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4996997 seqs, 3332346308 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 2 | AAQ97526 | Aaq97526 Human DNA |
| 2 | 3063 | 100.0 | 3063 | 4 | AAH76365 | Aah76365 Human PMS |
| 3 | 3063 | 100.0 | 3063 | 5 | AAH75042 | Aah75042 Nucleotid |
| 4 | 3063 | 100.0 | 3063 | 6 | ABK86089 | Abk86089 Human cDN |
| 5 | 3063 | 100.0 | 3063 | 6 | AAD39198 | Aad39198 Human mis |
| 6 | 3063 | 100.0 | 3063 | 6 | AAL48698 | Aal48698 Human mis |
| 7 | 3063 | 100.0 | 3063 | 6 | AAD45354 | Aad45354 Human MLH |
| 8 | 3063 | 100.0 | 3063 | 6 | AAD39770 | Aad39770 Human PMS |
| 9 | 3063 | 100.0 | 3063 | 8 | ABX12940 | Abx12940 DNA encod |
| 10 | 3063 | 100.0 | 3063 | 8 | ACA89704 | Aca89704 cDNA enco |
| 11 | 3063 | 100.0 | 3063 | 9 | ACD13762 | Acd13762 Human Mut |
| 12 | 3063 | 100.0 | 3063 | 9 | AAL57764 | Aal57764 Human mis |
| 13 | 3063 | 100.0 | 3063 | 9 | ADA06245 | Ada06245 DNA encod |
| 14 | 3063 | 100.0 | 3063 | 10 | ADC89608 | Adc89608 Human PMS |
| 15 | 3063 | 100.0 | 3063 | 10 | ADE85236 | Ade85236 Farnesyl |
| 16 | 3063 | 100.0 | 3063 | 10 | ADF17893 | Adf17893 Human PMS |
| 17 | 3063 | 100.0 | 3063 | 10 | ADG62892 | Adg62892 Human PMS |
| 18 | 3063 | 100.0 | 3063 | 10 | ADH62630 | Adh62630 Human mis |
| 19 | 3063 | 100.0 | 3063 | 10 | ADH60982 | Adh60982 Human cDN |
| 20 | 3063 | 100.0 | 3063 | 11 | ADI32109 | Adi32109 Human cDN |
| 21 | 3063 | 100.0 | 3063 | 12 | ADF78841 | Adf78841 Human mis |
| 22 | 3063 | 100.0 | 3063 | 12 | ADG46768 | Adg46768 Human MMR |
| 23 | 3063 | 100.0 | 3063 | 12 | ADO40067 | Ado40067 Human PMS |
| 24 | 3063 | 100.0 | 3063 | 12 | ADP66679 | Adp66679 Human mis |
| 25 | 3063 | 100.0 | 3063 | 13 | ADR13884 | Adr13884 Human DNA |
| 26 | 3063 | 100.0 | 3063 | 13 | ADT98687 | Adt98687 Human PMS |
| 27 | 3063 | 100.0 | 3063 | 13 | ADU77023 | Adu77023 Human mis |
| 28 | 3063 | 100.0 | 3063 | 13 | ADS84176 | Ads84176 Human lym |
| 29 | 3063 | 100.0 | 3063 | 14 | ADX58474 | Adx58474 Nucleotid |
| 30 | 3063 | 100.0 | 3063 | 14 | ADY53430 | Ady53430 Human PMS |
| 31 | 3012 | 98.3 | 3063 | 6 | AAD45433 | Aad45433 Human MLH |
| 32 | 3004 | 98.1 | 3239 | 14 | AEB22842 | Aeb22842 Human col |
| 33 | 2879 | 94.0 | 3121 | 14 | ADX08122 | Adx08122 Cyclin-de |
| 34 | 2110 | 68.9 | 2793 | 13 | ACN42335 | Acn42335 Human dia |
| 35 | 187 | 6.1 | 327 | 8 | ABX12507 | Abx12507 DNA repai |
| 36 | 187 | 6.1 | 327 | 9 | ADA06170 | Ada06170 Human EST |
| 37 | 154 | 5.0 | 559 | 5 | ABV48030 | Abv48030 Human pro |
| 38 | 126 | 4.1 | 478 | 9 | ACH23591 | Ach23591 Human adu |
| 39 | 115 | 3.8 | 184 | 3 | AAC22447 | Aac22447 Human sec |
| 40 | 96 | 3.1 | 189 | 2 | AAT19265 | Aat19265 Human gen |
| 41 | 82 | 2.7 | 534 | 5 | ABV18243 | Abv18243 Human pro |
| 42 | 60 | 2.0 | 60 | 6 | ABN46436 | Abn46436 Human spl |
| 43 | 50 | 1.6 | 50 | 6 | ABZ01968 | Abz01968 Human leu |
| 44 | 41 | 1.3 | 493 | 13 | ADQ79340 | Adq79340 Novel can |
| c 45 | 35 | 1.1 | 51 | 4 | AAL29142 | Aal29142 Human SNP |

ALIGNMENTS

RESULT 1
AAQ97526
ID AAQ97526 standard; cDNA; 3063 BP.
XX
AC AAQ97526;
XX
DT 27-MAR-1996 (first entry)
XX
DE Human DNA repair protein hMLH2 coding sequence.
XX
KW DNA repair protein; hMLH1; hMLH2; hMLH3; therapy; cancer; vectors;
KW DNA synthesis; diagnosis; disease; mutL4; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 81. .2873
FT /*tag= a
FT /product= "DNA repair protein."
FT /note= "The CDS may extend to position 2879 where a TGA
FT stop codon is located. There could possibly be an error
FT in the decoded protein in the specification since, if the
FT stop codon is the correct translation termination signal,
FT the C-terminal end of the protein should end with two
FT Threonine residues. It is possible that one of these has
FT been omitted."
XX
PN WO9520678-A1.
XX
PD 03-AUG-1995.
XX
PF 25-JAN-1995; 95WO-US001035.
XX
PR 27-JAN-1994; 94US-00187757.
PR 16-MAR-1994; 94US-00210143.
PR 23-AUG-1994; 94US-00294312.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Haseltine WA, Ruben SM, Wei Y, Adams MD, Fleischmann RD;
PI Fraser CM, Fuldner RA, Kirkness EF, Rosen CA;
XX

DR WPI; 1995-275461/36.
DR P-PSDB; AAR79009.
XX
PT Polynucleotide(s) encoding human mutL4 homologues, hMLH1, hMLH2 and hMLH3
PT - used for therapeutic treatment of, e.g. hereditary cancer.
XX
PS Claim 1; Fig 2; 124pp; English.
XX
CC The polynucleotides described in AAQ97525-27 encode the human analogues
CC of the prokaryotic mutL4 DNA repair gene. The polypeptides they encode
CC (AAR79008-R79010) are used for therapeutic purposes e.g. in the treatment
CC of cancer, esp. hereditary cancer. They may also be used for in vitro
CC manipulation of DNA, synthesis of DNA and the manufacture of DNA vectors
CC and in methods of diagnosing a disease or a susceptibilty to a disease
CC related to a mutation in the hMLH1, -2 or -3 DNA repair genes
XX
SQ Sequence 3063 BP; 1100 A; 503 C; 580 G; 880 T; 0 U; 0 Other;

Query Match 100.0%; Score 3063; DB 2; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
Db    1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy   61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
Db   61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy  121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
Db  121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy  181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
Db  181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy  241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Db  241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300

Qy  301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db  301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360

Qy  361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db  361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420

Qy  421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db  421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480

Qy  481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db  481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540

Qy  541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db  541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600

Qy  601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db  601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660

Qy  661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db  661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720

Qy  721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db  721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780

Qy  781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db  781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840

Qy  841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db  841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900

Qy  901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db  901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960

Qy  961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db  961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020

Qy 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080

Qy 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140

Qy 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200

Qy 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260

Qy 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320

Qy 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380

Qy 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440

Qy 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500

Qy 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560

Qy 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620

Qy 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680

Qy 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740

Qy 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800

Qy 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 14:43:11 ; Search time 7734 Seconds
(without alignments)
18529.717 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 23393541228 residues

Word size : 0

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 1804 | 58.9 | 1804 | 4 | CR610658 | CR610658 full-leng |
| | 2 | 1378 | 45.0 | 2365 | 4 | BC036376 | BC036376 Homo sapi |
| | 3 | 1294 | 42.2 | 2667 | 11 | DQ052952 | DQ052952 Homo sapi |
| | 4 | 844 | 27.6 | 3130 | 4 | CR859593 | CR859593 Pongo pyg |
| | 5 | 830 | 27.1 | 2667 | 11 | DQ052953 | DQ052953 Pan trogl |
| | 6 | 804 | 26.2 | 1011 | 5 | BX353664 | BX353664 BX353664 |
| | 7 | 780 | 25.5 | 780 | 7 | CN336501 | CN336501 170006001 |
| | 8 | 776 | 25.3 | 776 | 5 | BX117693 | BX117693 BX117693 |
| | 9 | 766 | 25.0 | 1021 | 3 | BM479838 | BM479838 AGENCOURT |
| | 10 | 711 | 23.2 | 733 | 7 | CN336498 | CN336498 170006001 |
| | 11 | 705 | 23.0 | 1117 | 3 | BM553209 | BM553209 AGENCOURT |
| c | 12 | 694 | 22.7 | 722 | 3 | BM677668 | BM677668 UI-E-EO1- |
| c | 13 | 677 | 22.1 | 704 | 5 | BQ771615 | BQ771615 UI-H-EZ1- |
| c | 14 | 634 | 20.7 | 638 | 1 | AL043809 | AL043809 DKFZp434P |
| c | 15 | 616 | 20.1 | 898 | 5 | BX435290 | BX435290 BX435290 |
| c | 16 | 615 | 20.1 | 746 | 5 | BU623174 | BU623174 UI-H-FL1- |
| | 17 | 604 | 19.7 | 903 | 5 | BX327629 | BX327629 BX327629 |
| | 18 | 590 | 19.3 | 641 | 6 | CB157373 | CB157373 K-EST0216 |
| c | 19 | 590 | 19.3 | 748 | 6 | CA415469 | CA415469 UI-H-EZ0- |
| | 20 | 579 | 18.9 | 668 | 8 | DN998411 | DN998411 TC106042 |
| c | 21 | 574 | 18.7 | 610 | 1 | AI636100 | AI636100 tz92g07.x |
| | 22 | 566 | 18.5 | 566 | 6 | CB161021 | CB161021 K-EST0220 |
| | 23 | 565 | 18.4 | 956 | 3 | BM800196 | BM800196 AGENCOURT |
| | 24 | 555 | 18.1 | 782 | 5 | BQ429685 | BQ429685 AGENCOURT |
| | 25 | 553 | 18.1 | 694 | 1 | AL705101 | AL705101 DKFZp686H |
| | 26 | 552 | 18.0 | 603 | 3 | BM723144 | BM723144 UI-E-EO1- |
| c | 27 | 528 | 17.2 | 623 | 2 | BE350913 | BE350913 ht63h02.x |
| | 28 | 525 | 17.1 | 558 | 6 | CB136057 | CB136057 K-EST0188 |
| c | 29 | 525 | 17.1 | 561 | 5 | BQ574496 | BQ574496 UI-H-EZ1- |
| c | 30 | 524 | 17.1 | 586 | 2 | BF056020 | BF056020 7k07e01.x |
| | 31 | 516 | 16.8 | 525 | 6 | CB268969 | CB268969 1007876 H |
| c | 32 | 513 | 16.7 | 706 | 6 | CB852984 | CB852984 UI-CF-FN0 |
| | 33 | 504 | 16.5 | 555 | 6 | CB131671 | CB131671 K-EST0181 |
| c | 34 | 503 | 16.4 | 893 | 5 | BX328949 | BX328949 BX328949 |
| c | 35 | 502 | 16.4 | 604 | 2 | BE350907 | BE350907 ht63g03.x |
| | 36 | 502 | 16.4 | 626 | 1 | AV730735 | AV730735 AV730735 |
| c | 37 | 486 | 15.9 | 500 | 1 | AI367805 | AI367805 qt56d07.x |
| | 38 | 486 | 15.9 | 4758 | 4 | CR749432 | CR749432 Homo sapi |
| c | 39 | 484 | 15.8 | 534 | 1 | AA573406 | AA573406 nm53d01.s |
| | 40 | 484 | 15.8 | 586 | 7 | CN336500 | CN336500 170004247 |
| c | 41 | 482 | 15.7 | 538 | 1 | AA573397 | AA573397 nm53c01.s |
| c | 42 | 481 | 15.7 | 524 | 1 | AI660351 | AI660351 we62b03.x |
| c | 43 | 479 | 15.6 | 492 | 1 | AI277404 | AI277404 qm58g01.x |
| | 44 | 477 | 15.6 | 820 | 5 | BX419980 | BX419980 BX419980 |
| | 45 | 477 | 15.6 | 835 | 2 | BG168340 | BG168340 602342315 |

ALIGNMENTS

RESULT 1
CR610658
LOCUS CR610658 1804 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DC010YK21 of Neuroblastoma Cot 25-normalized of Homo sapiens (human).
ACCESSION CR610658
VERSION CR610658.1 GI:50491465
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1804)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ InVitroGen Corporation 1600 Faraday Avenue
REFERENCE 2 (bases 1 to 1804)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES Location/Qualifiers
source 1. .1804
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC010YK21"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN

Query Match 58.9%; Score 1804; DB 4; Length 1804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1215 AATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATATGCATAATGATGAA 1274
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATATGCATAATGATGAA 60

Qy       1275 TCTGGAAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTGGTGACTTTGGTTAT 1334
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTGGTGACTTTGGTTAT 120

Qy       1335 GGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGAATGCATTTCAGGAC 1394
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGAATGCATTTCAGGAC 180

Qy       1395 ATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATAGTAAAACTTGTTTT 1454
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATAGTAAAACTTGTTTT 240

Qy       1455 ATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACCATATAGATGAGAGT 1514
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 ATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACCATATAGATGAGAGT 300

Qy       1515 GGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTTCTGCAGATGAGTGG 1574
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTTCTGCAGATGAGTGG 360

Qy       1575 AGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAACCTGTGAAAATTTTA 1634
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAACCTGTGAAAATTTTA 420

Qy       1635 GTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATCCAATCCCTGAACAA 1694
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATCCAATCCCTGAACAA 480

Qy       1695 ATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAGATAATAAATCTGGA 1754
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAGATAATAAATCTGGA 540

Qy       1755 AAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAACCCATGTCAGCAAGT 1814
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAACCCATGTCAGCAAGT 600

Qy       1815 GCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATCCTAAGACTAGTTTA 1874
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATCCTAAGACTAGTTTA 660

Qy       1875 GAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTGAAGAGGAAAAACTG 1934
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTGAAGAGGAAAAACTG 720

Qy       1935 AAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTCAAATGAAGAGAGCC 1994
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTCAAATGAAGAGAGCC 780

Qy       1995 ATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGATAAAACCCACCAGC 2054
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 ATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGATAAAACCCACCAGC 840

Qy       2055 GCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTAATCAACCAAAACTT 2114
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 GCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTAATCAACCAAAACTT 900

Qy       2115 GATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATATTAAAATGGTACAG 2174
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATATTAAAATGGTACAG 960

Qy       2175 ATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAAACAAAGTTGACTTA 2234
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 ATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAAACAAAGTTGACTTA 1020

Qy       2235 GAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTGATGCATGGCTAATG 2294
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 GAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTGATGCATGGCTAATG 1080

Qy       2295 ACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAGAAGCCCTGCTATTT 2354
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAGAAGCCCTGCTATTT 1140

Qy       2355 AAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAAAGCCAATTATGTTA 2414
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 AAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAAAGCCAATTATGTTA 1200

Qy       2415 ACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATAAAATGACAGCAGAT 2474
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 ACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATAAAATGACAGCAGAT 1260

Qy       2475 GACCAAAGATACAGTGGATCAACTTACCTGTCTGATCCTCGTCTTACAGCGAATGGTTTC 2534
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GACCAAAGATACAGTGGATCAACTTACCTGTCTGATCCTCGTCTTACAGCGAATGGTTTC 1320

Qy       2535 AAGATAAAATTGATACCAGGAGTTTCAATTACTGAAAATTACTTGGAAATAGAAGGAATG 2594
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AAGATAAAATTGATACCAGGAGTTTCAATTACTGAAAATTACTTGGAAATAGAAGGAATG 1380

Qy       2595 GCTAATTGTCTCCCATTCTATGGAGTAGCAGATTTAAAAGAAATTCTTAATGCTATATTA 2654
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 GCTAATTGTCTCCCATTCTATGGAGTAGCAGATTTAAAAGAAATTCTTAATGCTATATTA 1440

Qy       2655 AACAGAAATGCAAAGGAAGTTTATGAATGTAGACCTCGCAAAGTGATAAGTTATTTAGAG 2714
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AACAGAAATGCAAAGGAAGTTTATGAATGTAGACCTCGCAAAGTGATAAGTTATTTAGAG 1500

Qy       2715 GGAGAAGCAGTGCGTCTATCCAGACAATTACCCATGTACTTATCAAAAGAGGACATCCAA 2774
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GGAGAAGCAGTGCGTCTATCCAGACAATTACCCATGTACTTATCAAAAGAGGACATCCAA 1560

Qy       2775 GACATTATCTACAGAATGAAGCACCAGTTTGGAAATGAAATTAAAGAGTGTGTTCATGGT 2834
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 GACATTATCTACAGAATGAAGCACCAGTTTGGAAATGAAATTAAAGAGTGTGTTCATGGT 1620

Qy       2835 CGCCCATTTTTTCATCATTTAACCTATCTTCCAGAAACTACATGATTAAATATGTTTAAG 2894
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CGCCCATTTTTTCATCATTTAACCTATCTTCCAGAAACTACATGATTAAATATGTTTAAG 1680

Qy       2895 AAGATTAGTTACCATTGAAATTGGTTCTGTCATAAAACAGCATGAGTCTGGTTTTAAATT 2954
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 AAGATTAGTTACCATTGAAATTGGTTCTGTCATAAAACAGCATGAGTCTGGTTTTAAATT 1740

Qy       2955 ATCTTTGTATTATGTGTCACATGGTTATTTTTTAAATGAGGATTCACTGACTTGTTTTTA 3014
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 ATCTTTGTATTATGTGTCACATGGTTATTTTTTAAATGAGGATTCACTGACTTGTTTTTA 1800

Qy       3015 TATT 3018
             ||||
Db       1801 TATT 1804
```

RESULT 2
BC036376

```
LOCUS       BC036376               2365 bp    mRNA    linear   HTC 19-NOV-2003
DEFINITION  Homo sapiens PMS1 postmeiotic segregation increased 1 (S.
            cerevisiae), mRNA (cDNA clone IMAGE:4822417), containing
            frame-shift errors.
ACCESSION   BC036376
VERSION     BC036376.1  GI:22135671
KEYWORDS    HTC.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 2365)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 15:09:31 ; Search time 371 Seconds
(without alignments)
14675.664 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 3063 | 100.0 | 3063 | 3 | US-08-294-312B-3 | Sequence 3, Appli |
| | 2 | 3063 | 100.0 | 3063 | 3 | US-08-468-024B-3 | Sequence 3, Appli |
| | 3 | 3063 | 100.0 | 3063 | 3 | US-09-708-200-12 | Sequence 12, Appl |
| | 4 | 3063 | 100.0 | 3063 | 3 | US-09-023-655-1435 | Sequence 1435, Ap |
| | 5 | 3063 | 100.0 | 3063 | 3 | US-08-465-679-3 | Sequence 3, Appli |
| | 6 | 3063 | 100.0 | 3063 | 3 | US-09-788-657-8 | Sequence 8, Appli |
| | 7 | 3063 | 100.0 | 3063 | 3 | US-09-712-691-10 | Sequence 10, Appl |
| | 8 | 3063 | 100.0 | 3063 | 3 | US-09-707-468C-10 | Sequence 10, Appl |
| | 9 | 3063 | 100.0 | 3063 | 3 | US-10-641-068-8 | Sequence 8, Appli |
| | 10 | 2761 | 90.1 | 2817 | 3 | US-09-949-016-1916 | Sequence 1916, Ap |
| | 11 | 892 | 29.1 | 96845 | 3 | US-09-949-016-13658 | Sequence 13658, A |
| | 12 | 187 | 6.1 | 327 | 3 | US-08-187-757D-3 | Sequence 3, Appli |
| | 13 | 187 | 6.1 | 327 | 3 | US-08-210-143C-3 | Sequence 3, Appli |
| | 14 | 134 | 4.4 | 601 | 3 | US-09-949-016-65618 | Sequence 65618, A |
| | 15 | 115 | 3.8 | 184 | 3 | US-09-513-999C-26522 | Sequence 26522, A |
| | 16 | 50 | 1.6 | 50 | 3 | US-10-131-827-1959 | Sequence 1959, Ap |
| | 17 | 38 | 1.2 | 601 | 3 | US-09-949-016-65619 | Sequence 65619, A |
| c | 18 | 23 | 0.8 | 601 | 3 | US-09-949-016-151589 | Sequence 151589, |
| c | 19 | 23 | 0.8 | 601 | 3 | US-09-949-016-151591 | Sequence 151591, |
| | 20 | 23 | 0.8 | 143550 | 3 | US-09-949-016-14143 | Sequence 14143, A |
| c | 21 | 23 | 0.8 | 169334 | 3 | US-09-949-016-15999 | Sequence 15999, A |
| | 22 | 22 | 0.7 | 60 | 3 | US-08-294-312B-58 | Sequence 58, Appl |
| | 23 | 22 | 0.7 | 60 | 3 | US-08-468-024B-58 | Sequence 58, Appl |
| | 24 | 22 | 0.7 | 60 | 3 | US-08-465-679-58 | Sequence 58, Appl |
| c | 25 | 22 | 0.7 | 331814 | 3 | US-09-949-016-12008 | Sequence 12008, A |
| c | 26 | 22 | 0.7 | 331814 | 3 | US-09-949-016-17056 | Sequence 17056, A |
| c | 27 | 21 | 0.7 | 21 | 3 | US-08-294-312B-59 | Sequence 59, Appl |
| c | 28 | 21 | 0.7 | 21 | 3 | US-08-294-312B-61 | Sequence 61, Appl |
| | 29 | 21 | 0.7 | 21 | 3 | US-08-294-312B-70 | Sequence 70, Appl |
| | 30 | 21 | 0.7 | 21 | 3 | US-08-294-312B-72 | Sequence 72, Appl |
| c | 31 | 21 | 0.7 | 21 | 3 | US-08-468-024B-59 | Sequence 59, Appl |
| c | 32 | 21 | 0.7 | 21 | 3 | US-08-468-024B-61 | Sequence 61, Appl |
| | 33 | 21 | 0.7 | 21 | 3 | US-08-468-024B-70 | Sequence 70, Appl |
| | 34 | 21 | 0.7 | 21 | 3 | US-08-468-024B-72 | Sequence 72, Appl |
| c | 35 | 21 | 0.7 | 21 | 3 | US-08-465-679-59 | Sequence 59, Appl |
| c | 36 | 21 | 0.7 | 21 | 3 | US-08-465-679-61 | Sequence 61, Appl |
| | 37 | 21 | 0.7 | 21 | 3 | US-08-465-679-70 | Sequence 70, Appl |
| | 38 | 21 | 0.7 | 21 | 3 | US-08-465-679-72 | Sequence 72, Appl |
| | 39 | 21 | 0.7 | 58 | 3 | US-08-294-312B-60 | Sequence 60, Appl |
| | 40 | 21 | 0.7 | 58 | 3 | US-08-468-024B-60 | Sequence 60, Appl |
| | 41 | 21 | 0.7 | 58 | 3 | US-08-465-679-60 | Sequence 60, Appl |
| c | 42 | 21 | 0.7 | 601 | 3 | US-09-949-016-60541 | Sequence 60541, A |
| c | 43 | 21 | 0.7 | 32099 | 3 | US-09-949-016-16562 | Sequence 16562, A |
| | 44 | 21 | 0.7 | 59519 | 3 | US-09-949-016-13504 | Sequence 13504, A |
| c | 45 | 21 | 0.7 | 121970 | 3 | US-09-949-016-17216 | Sequence 17216, A |

ALIGNMENTS

RESULT 1
US-08-294-312B-3
; Sequence 3, Application US/08294312B
; Patent No. 6380369
; GENERAL INFORMATION:
; APPLICANT: Adams et al.
; TITLE OF INVENTION: Human DNA Mismatch Repair Proteins
; FILE REFERENCE: PF106P2
; CURRENT APPLICATION NUMBER: US/08/294,312B
; CURRENT FILING DATE: 1994-08-23
; PRIOR APPLICATION NUMBER: 08/210,143
; PRIOR FILING DATE: 1994-03-16
; PRIOR APPLICATION NUMBER: 08/187,757
; PRIOR FILING DATE: 1994-01-27
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3063
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (81)..(2879)
US-08-294-312B-3

Query Match 100.0%; Score 3063; DB 3; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Qy        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db        301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Qy        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db        361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Qy        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db        421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Qy        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db        481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Qy        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db        541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Qy        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db        601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Qy        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db        661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Qy        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db        721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Qy        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db        781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Qy        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db        841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Qy        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db        901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Qy        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db        961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Qy       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db       1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Qy       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db       1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Qy       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db       1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Qy       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db       1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Qy       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db       1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Qy       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db       1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
```

```
Qy       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db       1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Qy       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db       1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Qy       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db       1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Qy       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db       1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Qy       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db       1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Qy       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db       1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Qy       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db       1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Qy       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db       1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Qy       1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Db       1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Qy       1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Db       1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Qy       1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Db       1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Qy       2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Db       2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Qy       2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Db       2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Qy       2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Db       2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Qy       2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Db       2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Qy       2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Db       2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Qy       2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Db       2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Qy       2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460
Db       2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460
```

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 18:21:20 ; Search time 1567 Seconds
(without alignments)
16164.075 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 4134689005 residues

Word size : 0

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 3 | US-09-788-657-8 | Sequence 8, Appli |
| 2 | 3063 | 100.0 | 3063 | 3 | US-09-912-697-7 | Sequence 7, Appli |
| 3 | 3063 | 100.0 | 3063 | 3 | US-09-760-285-19 | Sequence 19, Appl |
| 4 | 3063 | 100.0 | 3063 | 5 | US-10-079-429-3 | Sequence 3, Appli |
| 5 | 3063 | 100.0 | 3063 | 6 | US-10-270-839-28 | Sequence 28, Appl |
| 6 | 3063 | 100.0 | 3063 | 6 | US-10-243-130-10 | Sequence 10, Appl |
| 7 | 3063 | 100.0 | 3063 | 6 | US-10-371-857-14 | Sequence 14, Appl |
| 8 | 3063 | 100.0 | 3063 | 6 | US-10-371-634-8 | Sequence 8, Appli |
| 9 | 3063 | 100.0 | 3063 | 6 | US-10-348-074-6 | Sequence 6, Appli |
| 10 | 3063 | 100.0 | 3063 | 6 | US-10-369-845-12 | Sequence 12, Appl |
| 11 | 3063 | 100.0 | 3063 | 7 | US-10-641-643-1435 | Sequence 1435, Ap |
| 12 | 3063 | 100.0 | 3063 | 7 | US-10-641-068-8 | Sequence 8, Appli |
| 13 | 3063 | 100.0 | 3063 | 7 | US-10-283-975A-453 | Sequence 453, App |
| 14 | 3063 | 100.0 | 3063 | 7 | US-10-813-502-10 | Sequence 10, Appl |
| 15 | 3063 | 100.0 | 3063 | 8 | US-10-714-228-1 | Sequence 1, Appli |
| 16 | 3063 | 100.0 | 3063 | 8 | US-10-850-370-10 | Sequence 10, Appl |
| 17 | 3063 | 100.0 | 3063 | 9 | US-10-933-034-1 | Sequence 1, Appli |
| 18 | 3063 | 100.0 | 3063 | 9 | US-10-901-650-10 | Sequence 10, Appl |
| 19 | 3063 | 100.0 | 3063 | 10 | US-11-056-776-14 | Sequence 14, Appl |
| 20 | 2799 | 91.4 | 2799 | 6 | US-10-109-791A-9 | Sequence 9, Appli |
| 21 | 560 | 18.3 | 611 | 4 | US-09-925-065A-562210 | Sequence 562210, |
| 22 | 233 | 7.6 | 419 | 7 | US-10-242-535A-25367 | Sequence 25367, A |
| 23 | 233 | 7.6 | 419 | 7 | US-10-085-783A-25367 | Sequence 25367, A |
| 24 | 154 | 5.0 | 559 | 8 | US-10-357-930-48049 | Sequence 48049, A |
| 25 | 126 | 4.1 | 478 | 3 | US-09-918-995-10803 | Sequence 10803, A |
| 26 | 82 | 2.7 | 534 | 8 | US-10-357-930-18234 | Sequence 18234, A |
| 27 | 60 | 2.0 | 60 | 3 | US-09-908-975-19184 | Sequence 19184, A |
| 28 | 50 | 1.6 | 50 | 6 | US-10-131-827-1959 | Sequence 1959, Ap |
| 29 | 48 | 1.6 | 120 | 7 | US-10-242-535A-21302 | Sequence 21302, A |
| 30 | 48 | 1.6 | 120 | 7 | US-10-085-783A-21302 | Sequence 21302, A |
| c 31 | 48 | 1.6 | 571 | 4 | US-09-925-065A-562206 | Sequence 562206, |
| c 32 | 48 | 1.6 | 622 | 4 | US-09-925-065A-562207 | Sequence 562207, |
| c 33 | 48 | 1.6 | 622 | 4 | US-09-925-065A-562208 | Sequence 562208, |
| c 34 | 25 | 0.8 | 25 | 9 | US-10-792-374-6 | Sequence 6, Appli |
| 35 | 25 | 0.8 | 25 | 10 | US-11-036-317-87094 | Sequence 87094, A |
| 36 | 25 | 0.8 | 25 | 10 | US-11-036-317-118726 | Sequence 118726, |
| 37 | 25 | 0.8 | 25 | 10 | US-11-036-317-175306 | Sequence 175306, |
| 38 | 25 | 0.8 | 25 | 10 | US-11-036-317-209992 | Sequence 209992, |
| 39 | 25 | 0.8 | 25 | 10 | US-11-036-317-222017 | Sequence 222017, |
| 40 | 25 | 0.8 | 25 | 10 | US-11-036-317-233940 | Sequence 233940, |
| 41 | 25 | 0.8 | 25 | 10 | US-11-036-317-281574 | Sequence 281574, |
| 42 | 25 | 0.8 | 25 | 10 | US-11-036-317-368945 | Sequence 368945, |
| 43 | 23 | 0.8 | 478 | 7 | US-10-702-075-254 | Sequence 254, App |
| 44 | 22 | 0.7 | 60 | 5 | US-10-079-429-58 | Sequence 58, Appl |
| c 45 | 22 | 0.7 | 575 | 4 | US-09-925-065A-272278 | Sequence 272278, |

ALIGNMENTS

RESULT 1
US-09-788-657-8
; Sequence 8, Application US/09788657
; Patent No. US20020123149A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/09/788,657
; CURRENT FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 3063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-788-657-8

Query Match 100.0%; Score 3063; DB 3; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

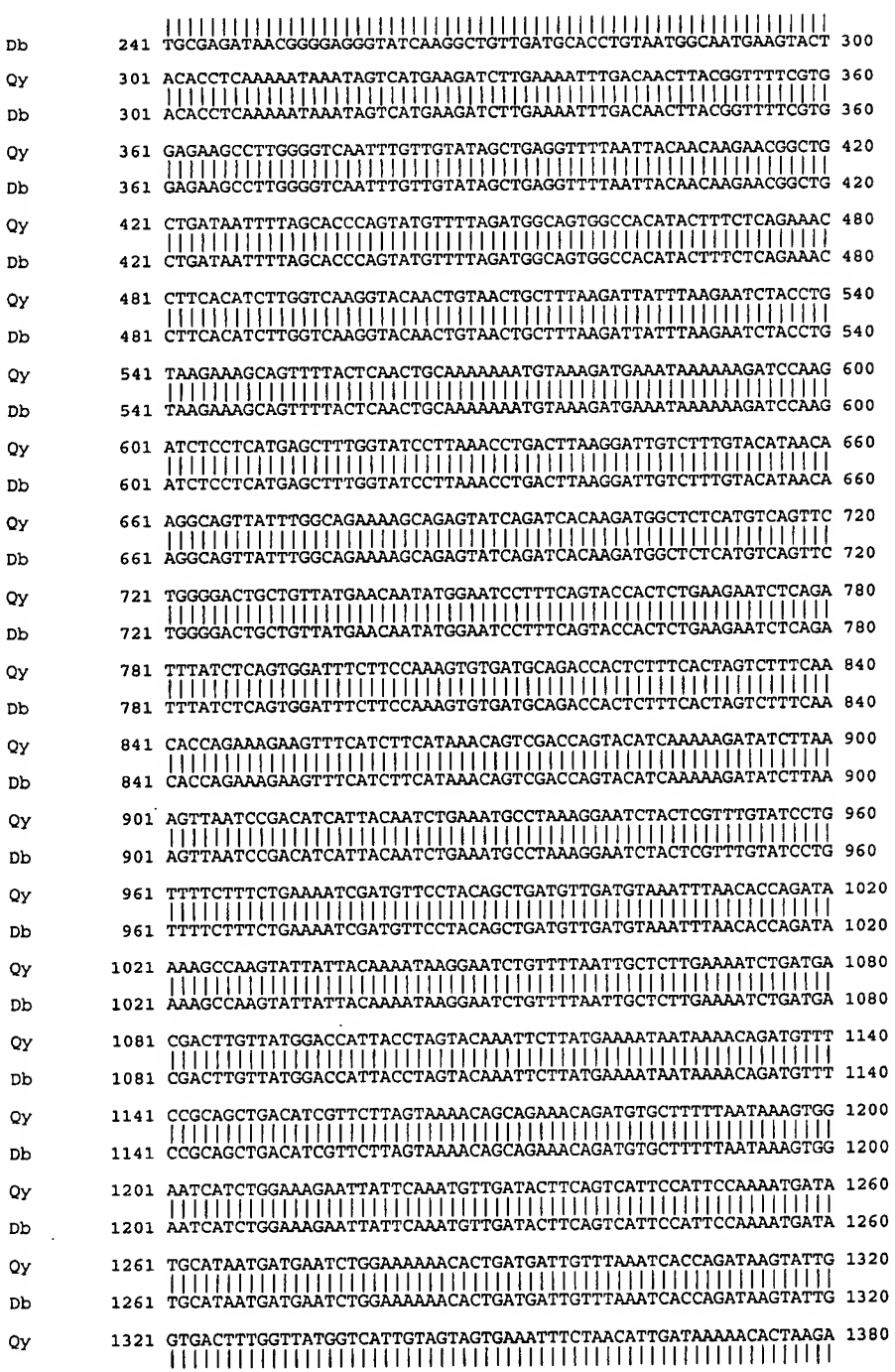
```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240

Qy        241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
```


Db 241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Qy 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Qy 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Qy 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Qy 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Qy 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Qy 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Qy 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Qy 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Qy 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Qy 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Qy 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Qy 961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db 961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Qy 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Qy 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Qy 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Qy 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Qy 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Qy 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Qy 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Qy 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Qy 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Qy 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Qy 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Qy 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Qy 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Qy 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Qy 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Db 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Qy 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Db 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Qy 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Db 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Qy 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Db 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Qy 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Db 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Qy 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Db 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Qy 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Db 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Qy 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Db 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Qy 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Db 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Qy 2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460
Db 2401 AGCCAATTATGTTAACAGAGAGTCTTTTTAATGGATCTCATTATTTAGACGTTTTATATA 2460

GenCore version 5.1.6


OM nucleic - nucleic search, using sw model

Run on: December 26, 2005, 20:30:17 ; Search time 244 Seconds
(without alignments)
6514.585 Million cell updates/sec

Title: US-10-079-429A-3
Perfect score: 3063
Sequence: 1 ggcacgagtggctgcttgcg.....aacgtaaataaactaataac 3063

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4168288 seqs, 259477437 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8336576

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 3063 | 100.0 | 3063 | 7 | US-11-188-743-8 | Sequence 8, Appli |
| 2 | 25 | 0.8 | 25 | 7 | US-11-121-849-235011 | Sequence 235011, |
| 3 | 25 | 0.8 | 25 | 7 | US-11-121-849-235012 | Sequence 235012, |
| 4 | 25 | 0.8 | 25 | 7 | US-11-121-849-235013 | Sequence 235013, |
| 5 | 25 | 0.8 | 25 | 7 | US-11-121-849-235014 | Sequence 235014, |
| 6 | 25 | 0.8 | 25 | 7 | US-11-121-849-235015 | Sequence 235015, |
| 7 | 25 | 0.8 | 25 | 7 | US-11-121-849-235016 | Sequence 235016, |
| 8 | 25 | 0.8 | 25 | 7 | US-11-121-849-235017 | Sequence 235017, |
| 9 | 25 | 0.8 | 25 | 7 | US-11-121-849-235018 | Sequence 235018, |
| 10 | 25 | 0.8 | 25 | 7 | US-11-121-849-235019 | Sequence 235019, |
| 11 | 25 | 0.8 | 25 | 7 | US-11-121-849-235020 | Sequence 235020, |
| 12 | 25 | 0.8 | 25 | 7 | US-11-121-849-235021 | Sequence 235021, |
| c 13 | 21 | 0.7 | 2007 | 6 | US-10-750-185-28749 | Sequence 28749, A |
| c 14 | 20 | 0.7 | 201 | 6 | US-10-995-561-78264 | Sequence 78264, A |
| c 15 | 20 | 0.7 | 201 | 6 | US-10-995-561-78285 | Sequence 78285, A |
| c 16 | 20 | 0.7 | 201 | 6 | US-10-995-561-78344 | Sequence 78344, A |
| c 17 | 20 | 0.7 | 201 | 6 | US-10-995-561-78367 | Sequence 78367, A |
| 18 | 20 | 0.7 | 1593 | 6 | US-10-750-185-31134 | Sequence 31134, A |
| 19 | 20 | 0.7 | 3238 | 6 | US-10-750-185-53560 | Sequence 53560, A |
| c 20 | 20 | 0.7 | 53331 | 6 | US-10-995-561-13476 | Sequence 13476, A |
| 21 | 20 | 0.7 | 177175 | 7 | US-11-121-086-79 | Sequence 79, Appl |
| 22 | 19 | 0.6 | 19 | 8 | US-11-101-244-382080 | Sequence 382080, |
| 23 | 19 | 0.6 | 19 | 8 | US-11-101-244-382081 | Sequence 382081, |
| 24 | 19 | 0.6 | 19 | 8 | US-11-101-244-382082 | Sequence 382082, |
| 25 | 19 | 0.6 | 19 | 8 | US-11-101-244-382083 | Sequence 382083, |
| 26 | 19 | 0.6 | 19 | 8 | US-11-101-244-382084 | Sequence 382084, |
| 27 | 19 | 0.6 | 19 | 8 | US-11-101-244-382085 | Sequence 382085, |
| 28 | 19 | 0.6 | 19 | 8 | US-11-101-244-382086 | Sequence 382086, |
| 29 | 19 | 0.6 | 19 | 8 | US-11-101-244-382087 | Sequence 382087, |
| 30 | 19 | 0.6 | 19 | 8 | US-11-101-244-382088 | Sequence 382088, |
| 31 | 19 | 0.6 | 19 | 8 | US-11-101-244-382089 | Sequence 382089, |
| 32 | 19 | 0.6 | 19 | 8 | US-11-101-244-382090 | Sequence 382090, |
| 33 | 19 | 0.6 | 19 | 8 | US-11-101-244-382091 | Sequence 382091, |
| 34 | 19 | 0.6 | 19 | 8 | US-11-101-244-382092 | Sequence 382092, |
| 35 | 19 | 0.6 | 19 | 8 | US-11-101-244-382093 | Sequence 382093, |
| 36 | 19 | 0.6 | 19 | 8 | US-11-101-244-382094 | Sequence 382094, |
| 37 | 19 | 0.6 | 19 | 8 | US-11-101-244-382095 | Sequence 382095, |
| 38 | 19 | 0.6 | 19 | 8 | US-11-101-244-382096 | Sequence 382096, |
| 39 | 19 | 0.6 | 19 | 8 | US-11-101-244-382097 | Sequence 382097, |
| 40 | 19 | 0.6 | 19 | 8 | US-11-101-244-382098 | Sequence 382098, |
| 41 | 19 | 0.6 | 19 | 8 | US-11-101-244-382099 | Sequence 382099, |
| 42 | 19 | 0.6 | 19 | 8 | US-11-101-244-382100 | Sequence 382100, |
| 43 | 19 | 0.6 | 19 | 8 | US-11-101-244-382101 | Sequence 382101, |
| 44 | 19 | 0.6 | 19 | 8 | US-11-101-244-382102 | Sequence 382102, |
| 45 | 19 | 0.6 | 19 | 8 | US-11-101-244-382103 | Sequence 382103, |

ALIGNMENTS

RESULT 1
US-11-188-743-8
; Sequence 8, Application US/11188743
; Publication No. US20050272140A1
; GENERAL INFORMATION:
; APPLICANT: Nicolaides, Nicholas
; APPLICANT: Sass, Philip
; APPLICANT: Kinzler, Kenneth
; APPLICANT: Grasso, Luigi
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: Methods for generating hypermutable
; TITLE OF INVENTION: yeast
; FILE REFERENCE: 01107.00097
; CURRENT APPLICATION NUMBER: US/11/188,743
; CURRENT FILING DATE: 2005-07-26
; PRIOR APPLICATION NUMBER: US/10/641,068
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: US/09/788,657
; PRIOR FILING DATE: 2001-02-21
; PRIOR APPLICATION NUMBER: 60/184,336
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 3063
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-188-743-8

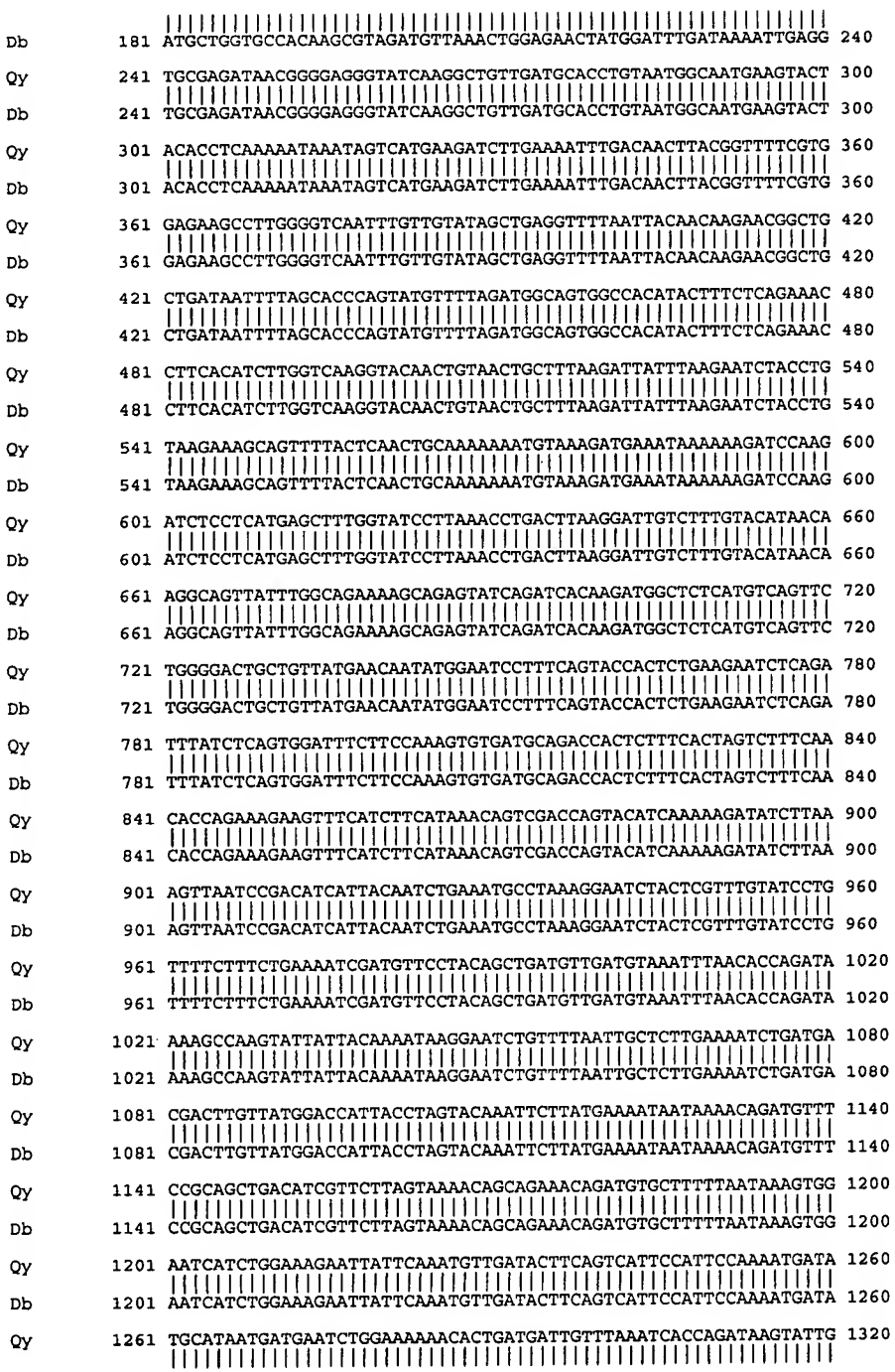
Query Match 100.0%; Score 3063; DB 7; Length 3063;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGCACGAGTGGCTGCTTGCGGCTAGTGGATGGTAATTGCCTGCCTCGCGCTAGCAGCAAG 60

Qy         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCTCTGTTAAAAGCGAAAATGAAACAATTGCCTGCGGCAACAGTTCGACTCCTTTCAA 120

Qy        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GTTCTCAGATCATCACTTCGGTGGTCAGTGTTGTAAAAGAGCTTATTGAAAACTCCTTGG 180

Qy        181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
```


Db 181 ATGCTGGTGCCACAAGCGTAGATGTTAAACTGGAGAACTATGGATTTGATAAAATTGAGG 240
Qy 241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Db 241 TGCGAGATAACGGGGAGGGTATCAAGGCTGTTGATGCACCTGTAATGGCAATGAAGTACT 300
Qy 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Db 301 ACACCTCAAAAATAAATAGTCATGAAGATCTTGAAAATTTGACAACTTACGGTTTTCGTG 360
Qy 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Db 361 GAGAAGCCTTGGGGTCAATTTGTTGTATAGCTGAGGTTTTAATTACAACAAGAACGGCTG 420
Qy 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Db 421 CTGATAATTTTAGCACCCAGTATGTTTTAGATGGCAGTGGCCACATACTTTCTCAGAAAC 480
Qy 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Db 481 CTTCACATCTTGGTCAAGGTACAACTGTAACTGCTTTAAGATTATTTAAGAATCTACCTG 540
Qy 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Db 541 TAAGAAAGCAGTTTTACTCAACTGCAAAAAAATGTAAAGATGAAATAAAAAAGATCCAAG 600
Qy 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Db 601 ATCTCCTCATGAGCTTTGGTATCCTTAAACCTGACTTAAGGATTGTCTTTGTACATAACA 660
Qy 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Db 661 AGGCAGTTATTTGGCAGAAAAGCAGAGTATCAGATCACAAGATGGCTCTCATGTCAGTTC 720
Qy 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Db 721 TGGGGACTGCTGTTATGAACAATATGGAATCCTTTCAGTACCACTCTGAAGAATCTCAGA 780
Qy 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Db 781 TTTATCTCAGTGGATTTCTTCCAAAGTGTGATGCAGACCACTCTTTCACTAGTCTTTCAA 840
Qy 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Db 841 CACCAGAAAGAAGTTTCATCTTCATAAACAGTCGACCAGTACATCAAAAAGATATCTTAA 900
Qy 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Db 901 AGTTAATCCGACATCATTACAATCTGAAATGCCTAAAGGAATCTACTCGTTTGTATCCTG 960
Qy 961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Db 961 TTTTCTTTCTGAAAATCGATGTTCCTACAGCTGATGTTGATGTAAATTTAACACCAGATA 1020
Qy 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Db 1021 AAAGCCAAGTATTATTACAAAATAAGGAATCTGTTTTAATTGCTCTTGAAAATCTGATGA 1080
Qy 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Db 1081 CGACTTGTTATGGACCATTACCTAGTACAAATTCTTATGAAAATAATAAAACAGATGTTT 1140
Qy 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Db 1141 CCGCAGCTGACATCGTTCTTAGTAAAACAGCAGAAACAGATGTGCTTTTTAATAAAGTGG 1200
Qy 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Db 1201 AATCATCTGGAAAGAATTATTCAAATGTTGATACTTCAGTCATTCCATTCCAAAATGATA 1260
Qy 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Db 1261 TGCATAATGATGAATCTGGAAAAAACACTGATGATTGTTTAAATCACCAGATAAGTATTG 1320
Qy 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Db 1321 GTGACTTTGGTTATGGTCATTGTAGTAGTGAAATTTCTAACATTGATAAAAACACTAAGA 1380
Qy 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Db 1381 ATGCATTTCAGGACATTTCAATGAGTAATGTATCATGGGAGAACTCTCAGACGGAATATA 1440
Qy 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Db 1441 GTAAAACTTGTTTTATAAGTTCCGTTAAGCACACCCAGTCAGAAAATGGCAATAAAGACC 1500
Qy 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Db 1501 ATATAGATGAGAGTGGGGAAAATGAGGAAGAAGCAGGTCTTGAAAACTCTTCGGAAATTT 1560
Qy 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Db 1561 CTGCAGATGAGTGGAGCAGGGGAAATATACTTAAAAATTCAGTGGGAGAGAATATTGAAC 1620
Qy 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Db 1621 CTGTGAAAATTTTAGTGCCTGAAAAAAGTTTACCATGTAAAGTAAGTAATAATAATTATC 1680
Qy 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Db 1681 CAATCCCTGAACAAATGAATCTTAATGAAGATTCATGTAACAAAAAATCAAATGTAATAG 1740
Qy 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Db 1741 ATAATAAATCTGGAAAAGTTACAGCTTATGATTTACTTAGCAATCGAGTAATCAAGAAAC 1800
Qy 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Db 1801 CCATGTCAGCAAGTGCTCTTTTTGTTCAAGATCATCGTCCTCAGTTTCTCATAGAAAATC 1860
Qy 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Db 1861 CTAAGACTAGTTTAGAGGATGCAACACTACAAATTGAAGAACTGTGGAAGACATTGAGTG 1920
Qy 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Db 1921 AAGAGGAAAAACTGAAATATGAAGAGAAGGCTACTAAAGACTTGGAACGATACAATAGTC 1980
Qy 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Db 1981 AAATGAAGAGAGCCATTGAACAGGAGTCACAAATGTCACTAAAAGATGGCAGAAAAAAGA 2040
Qy 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Db 2041 TAAAACCCACCAGCGCATGGAATTTGGCCCAGAAGCACAAGTTAAAAACCTCATTATCTA 2100
Qy 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Db 2101 ATCAACCAAAACTTGATGAACTCCTTCAGTCCCAAATTGAAAAAAGAAGGAGTCAAAATA 2160
Qy 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Db 2161 TTAAAATGGTACAGATCCCCTTTTCTATGAAAAACTTAAAAATAAATTTTAAGAAACAAA 2220
Qy 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Db 2221 ACAAAGTTGACTTAGAAGAGAAGGATGAACCTTGCTTGATCCACAATCTCAGGTTTCCTG 2280
Qy 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Db 2281 ATGCATGGCTAATGACATCCAAAACAGAGGTAATGTTATTAAATCCATATAGAGTAGAAG 2340
Qy 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400
Db 2341 AAGCCCTGCTATTTAAAAGACTTCTTGAGAATCATAAACTTCCTGCAGAGCCACTGGAAA 2400